

25<sup>th</sup> Meeting of the Wiesbaden Group on Business Registers  
- International Roundtable on Business Survey Frames

Tokyo, 8 – 11 November 2016

*Claude Macchi*  
*Swiss Federal Statistical Office (FSO)*  
*Session No. 6*

*Business Demography and Data Products from the Business Register*

***The Swiss SBR, administrative data and the production of business statistic data***

## **1. Introduction**

Driven by the need to reduce production costs and to decrease the burden on respondents and given the increased need of information, the Swiss Federal Statistical Office (FSO) is moving from a traditional, questionnaire-based production of business statistics to a centralised data system, based on the statistical business register (SBR) as well as on various administrative data, which are also used for the estimation and plausibilisation of data.

The challenge in this context was to develop a centralised system for the preparation of individual data based on the content of the SBR and on administrative data, for the production and analysis of business statistics. After three years of intense efforts the first statistic on enterprises (STATENT) was published mid-November 2013, bringing to an end not only the work of conceptualisation and integration but also demonstrating the viability of the concepts in concrete terms.

Like a jigsaw puzzle, the modernisation of the business statistics involved the collection and integration of data from the SBR, the National Buildings and Dwellings Register (RBD), the Register of Business Identification Numbers (IDE), the Profiling, cantonal agricultural registers, the Old Age and Survivors Insurance (OASI), the Federal Tax Administration as well as surveys for updating the SBR, the National Classification of Economic Activities (NOGA), and finally the Employment Statistics (ES), to create a picture that would be qualitatively coherent and usable for statistical production.

After this first phase, the development of the centralised system continued with the use of new administrative sources and the additional production of integrated statistical data. The newly integrated customs declarations will provide information, among others, on the imported and exported goods, their value, their origin and destination etc. and thus make possible, in combination with other data stored into the centralized data system, the production of the Foreign Affiliate Statistics (FATS). For the first time this year the FSO is also going to publish the business demography statistics based on the SBR and administrative data. *Sine qua non* condition for this evolution is the unavoidable harmonisation of metadata through all the concerned data as well as the widespread implementation of unique identifiers for enterprises.

## **2. A new approach for business statistics**

A basic step for the construction of a centralised production system for register-based statistics is the promotion of the harmonisation of data and in this way the creation of a solid basis for comparing national and international statistical information, which leads to a significant strengthening of their use potential. The production of the first STATENT data in 2013 was the outcome of a conceptualisation and standardisation work which demonstrates the viability of the concepts in concrete terms. The challenge was not to define new questions, or even to prepare questionnaires and organise the logistics, but rather to develop the basis for analysis and integration of the data contained in registers and coming from administrative sources for statistical purposes. These years have been devoted to preparing, transforming, adapting and standardising business data so that it can be used for statistical production.

## **3. The legal frame**

The availability as well as the implementation of administrative data into the statistical production processes of the FSO needed a legal base. This lack was covered in January 2008, as a new article on the “Principles of data collection” of the Swiss Federal Statistical Act (FStatA) came into force. It states that

*“Insofar as the Confederation has the necessary data available or that such data becomes available to an organisation subject to this Act in the implementation of federal law (federal administrative data), separate surveys need not be conducted by federal statistical bodies (direct surveys, indirect surveys or surveys based on observations and measurements).”<sup>1</sup>*

This should have guaranteed FSO access to the administrative data produced within the frame of the central national administration. If this was indeed the case for some administrative data (OASI- and VAT data first and customs data more recently), there are other data, like tax data, which are actually still unreachable. The acquisition of administrative data in Switzerland unfortunately is often related to long and tiresome discussions with the correspondent ministries.

The reasons for this bottleneck are various: diffidence concerning the observance of the data protection laws by the statistics producers, fear of loss of confidence in the administration (people send e.g. sensitive data to the tax administration which are then made available to other ministries), loss of exclusivity on certain information as well as reluctance in introducing changes etc.

## **4. A widespread use of administrative data in the statistical production**

A widespread use of administrative data goes hand in hand with the definition of framework conditions, with the implementation of unique identifiers, the use of standardised metadata as well as common tools for data administration, transformation and exchange.

---

<sup>1</sup> Federal Statistics Act (FStatA), RS 431.01, of 9 October 1992; <https://www.admin.ch/opc/en/classified-compilation/19920252/index.html>

In January 2011 came into force in Switzerland the regulation on the unique identifier for businesses (UIDV)<sup>2</sup> which entrusted the FSO with the construction and the administration of a central register of unique identifiers for all enterprises registered in the SBR, company registers, social security registers, tax and VAT registers, branch registers as well as in customs registers. After a 5 years implementation period, the administrative registers listed in the mentioned law have now all introduced the unique identifier (UID), simplifying thus on the one hand a countrywide standardised administrative handling of the businesses and allowing on the other hand the use of administrative data of these enterprises for statistical purposes.

This first fundamental step towards the use of administrative data by the FSO was realised in parallel with other measures in order to create a common denominator, a harmonised base for the production of business statistics. The data coming from the VAT registers, the OASI-funds and the customs register, already identified by the UID and therefore also linked with the SBR, were linked with the standardised metadata (classifications, variables etc.) stored in the central FSO metadata registry.

Administrative data are not collected and limited for only one specific statistic, but are made available for the whole statistical production, reused in more statistics, according to the “produce once and use many times” principle. Data and the related metadata, linked with the SBR through the UID, are centrally stored in a joint platform and made accessible to all statistical producers. This also facilitates and promotes the integration and dissemination of new administrative data sets in the statistical office. In this way, the FSO did the first step towards a centralised data system, laying the base for a production of harmonised, coherent and comparable statistics and thus doing an important groundwork for the construction of an integrated production system for business statistics.

## **5. The production of the STATENT data, or from a paper-based to a register-based business census**

Launched in 2008, developed further in 2011-12, and finalised in 2013, the modernisation of the federal business census was the second major undertaking aimed at modernising the Swiss official statistics following that of the census of the Swiss population completed in 2010. The main objective of this complex and ambitious project was to eliminate the need to send out 500,000 questionnaires to businesses every three or four years, by making use of register and administrative data and thus simplifying and reducing the burden for companies. Since 2013 the FSO is now producing yearly individual STATENT data built with data coming from registers (SBR, agricultural registers etc.), administrative sources (principally OASI data, but also with inputs coming from VAT and customs declarations), profiling activities as well as a few targeted surveys (e.g. the business enquiry ERST). STATENT data are first of all the springboard for the production of statistics on the structure of the businesses as well as the newly revised business demography, which is going to be published this year in October for the first time. These same data are becoming a fundamental source for more and more statistics produced by the FSO, such as cross-border statistics, apprentice statistics, bankruptcy statistics etc. and via the analysis of commuting patterns the definition of Swiss metropolitan areas.

---

<sup>2</sup> Verordnung über die Unternehmens-Identifikationsnummer (UIDV), RS 431.031, of 26 January 2011; <https://www.admin.ch/opc/de/classified-compilation/20101476/index.html> (this regulation is not available in English)

## 6. A centralised statistical data system as base for an integrated system for the production of business statistics

The system we are aiming for at the FSO is structured in two parts: a first block, a centralised statistical data system, focused on the collection and preparation of data, built around the SBR, administrative data as well as profiling activities, and a second block, where the business statistics are produced and analysed, based on the data prepared and made available by the data system.

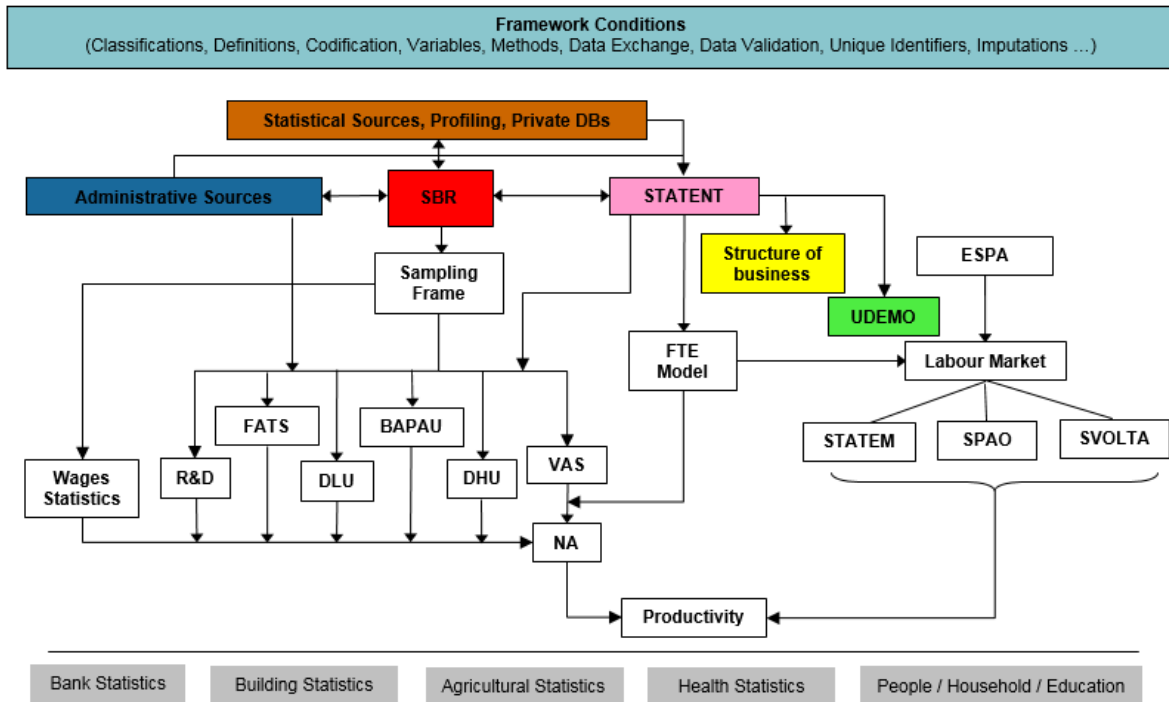


Figure 1: Integrated system for the production of business statistics at the FSO

As mentioned above, the basic elements of this data system are the framework conditions. The use of standardised concepts, classifications, variables, metadata and statistical methodology are necessary preconditions for both the centralised data system as well as the integrated production system. But all this cannot work without a clear strategy, strongly supported by the hierarchy, an appropriate organisational structure as well as an overall strong centralised coordination.

The data system is at the moment composed by five basic data sources: the SBR, profiling activities, administrative sources, statistical surveys and external data providers. Efforts are made by using administrative data as much as possible in substituting data collected through surveys, even if there is the awareness that surveys cannot be fully substituted in short and medium term and that at the moment a lot of statistical producers still collect data for their specific projects.

Efforts are in any case constantly made to enhance the questionnaires which are sent to the businesses. Special emphasis is put on the reduction of the time needed to fill in the questionnaires and synergies created between different surveys, for example by dropping same or similar questions and avoiding to send several questionnaires to the same businesses in a short lapse of time. In parallel we are also

strengthening the profiling activities, supporting above all direct contact to medium and large businesses, especially in the context of multinational enterprises.

First reflections about the use of Big Data have also been launched, but these discussions are actually in an initial phase and still far away from a regular implementation of this source into the production of statistics.

## **7. Advantages and problems in the use of administrative data for the production of statistics**

The use of administrative data entails several benefits, both from a statistical and from an organisational point of view, but it also leads to inevitable adaptations in the production processes, with ruptures in statistical time series and organisational reorganisations.

With the use of administrative data the amount of questionnaires sent out by the FSO decreased hugely. The population census was realised in 2010 without questionnaires for the first time, the census of enterprises was carried out for the last time in 2008. The suppression of questionnaires enabled us to reduce both our production costs and respondent burden substantially which is line with demands from enterprises, the national economic organisations and government.

At the same time, the introduction of register and administrative data widened the data sources for official statistics, allowing the creation of new products and the redesign of already existing statistics. Besides other advantages, one of the fundamental asset of administrative data is their high coverage of the observed population, which allows also a completion of the register content. Lots of information, which could not have been collected by surveys, are gathered through data coming from public administrations. For example, thanks to the use of OASI registers, which contain information on a huge number of small enterprises until now unknown to the SBR, the amount of businesses in Switzerland rose by more than 45% in the first year of the STATENT and the birth of enterprises increased by 275% in the first publication of the business demography based on these administrative data.

On the other hand, we should not underestimate the fact that administrative data are not produced for statistical purposes and have therefore to be processed and adapted, in order to make them usable for the statistical production and to guarantee comparability and coherence of data. This processes which is carried out in the centralized data system, asks for an inevitable adaptation of the received data to the statistical IT environment, plausibility checks and improvements, matching and imputation of data etc.

The use of administrative data also entails open questions concerning the production schedule of the statistics. Due to the fact that these data are usually supplied in a higher cadence than survey data, questions arise concerning the update and the improvement of certain individual data made available through the centralised data system. This could require a quite intense production of new data and time series, making longitudinal series analysis difficult, with important consequences for the production of statistics within the integrated system.

Should for example codes of economic activities in the data which are made available for the production of statistics be updated on the spot (event several times a year) or should they be frozen for a defined period (e.g. two, three or more years) in order to guarantee a stability of the statistics to be produced?