6. September 2016

25th Meeting of the Wiesbaden Group on Business Registers - International Roundtable on Business Survey Frames
Tokyo, 8 – 11 November 2016 *Roland Sturm* *Federal Statistical Office, Germany* *Session No. 2* *Role of Business Registers*
Users as producers – integrating survey information in the Statistical Business Register

1 Introduction

Modern business statistics in many statistical systems rely on a statistical business register (SBR) as an infrastructural, conceptual and methodological backbone. This puts the statistical business register in a most responsible position:

- It shall serve as frame for statistical surveys (for coherent sampling and extrapolation).

- It shall be a tool to acquire, combine and connect existing data in the statistical office as well as in other institutions, especially in public administration (i.a. to minimise response burden).

- It shall be an assisting instrument for the statisticians in every phase of the production process of business statistics.

To comply with these demands, the SBR must contain a set of population of units of different character and their inter-relations: The SBR shall contain all relevant reporting units (units providing information to the data collector), observation units (units about which information is provided) and statistical units (units a statistical output refers to) and all relations between these units.

Some years ago the German SBR contained only very limited information about the surveys which used the SBR and about the survey populations. This information consisted of codes indicating which register units were chosen as respondents for which surveys.

The new database for the German SBR which has been introduced mid-2014 detailed the data about survey participants widely. In line with this extension of the data base the survey departments in the Statistical Offices have been attributed an active role in the process of the register maintenance.

An additional layer has been introduced in the data architecture of the German SBR by the end of year 2015: Now survey departments have the possibility not only to store detailed information about their survey populations, but to introduce and maintain survey specific subunits of the units they find in the core SBR.

2. German Statistical Business Register in the 1990s

In the early 1990s, German official statistics built up an SBR that covered all economic sectors, taking stock from several existing sector-specific registers that had been run by subject matter departments till that date. The technical data base of one of these sector registers served as blueprint for the SBR data base. In this way German official statistics managed to build up a comprehensive SBR within short time. The German Business Register System, as it had been named, consisted of 16 regional data bases and a central annual copy of them. The SBR was maintained and updated by regular deliveries of administrative files and by information passed on by survey departments. Concerning the business surveys in Germany, the situation was – and to some extend still is – as follows:

- The German statistical system consists of a large number of branch (stove pipe) statistical surveys. In line with the general principle of federal administration in Germany most of the statistical surveys are conducted in a decentralised way in the (meanwhile) 14 statistical offices of the Länder. About 100 business surveys are register based, about 75 of them are conducted in decentralised manner by the regional offices.

- Practical working arrangements between the register and the survey departments may diverge in the statistical offices due to the autonomy offered by the existence of the regional SBR data bases and the decentralised way of survey conduction.

In consequence there emerged tendencies in the survey departments to "move out" from the SBR or – less rigorous – reluctance to "give back" to the SBR information obtained by survey conduction. Both issues became relevant obstacles for the use and running the SBR in the desired efficient way. As severe shortcoming had been identified the very restricted possibilities of the survey departments to document their populations in the SBR data base and lacking procedures for consistent and coherent participation of the survey statisticians in the SBR maintenance. The old SBR system contained basically "survey identifiers" – codes for each survey – and stored these as characteristics without time reference to register units in the core SBR section. For information on earlier survey participations one had to look in older backup copies of the SBR. The reference year was not listed so that the correct assignment of the reference year was to some extent guesswork or depended on in-depth knowledge about the surveys and the register maintenance procedures.

Figure 1:
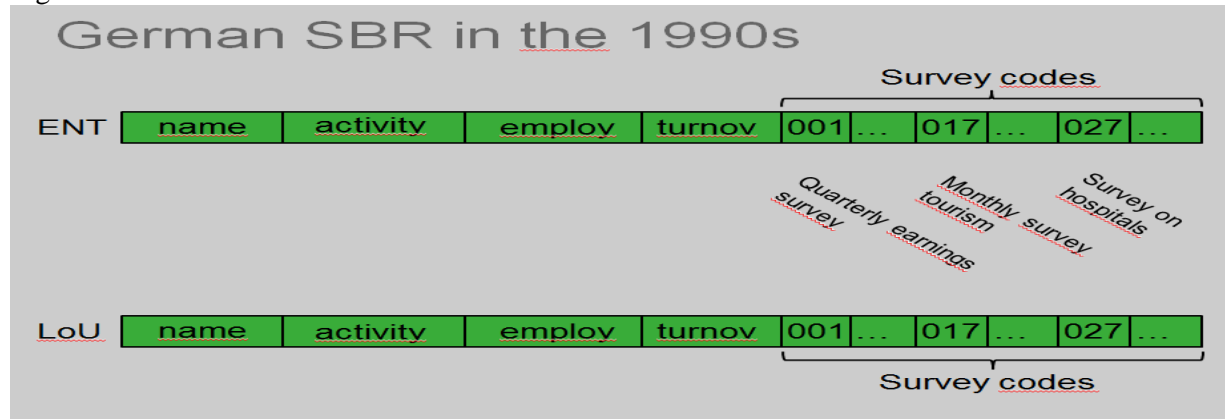
3. Issues of register maintenance and how to handle them

Commonly – even if not mentioned explicitly – there is an underlying idea of „producers" of the SBR and of „consumers" of register information. This simplifying view tends to imply that there are some whose task it is to provide a register and some who make use of it. Such a simplified attribution of roles runs the risk of suboptimal (or inadequate) quality and unsatisfactory use of the SBR. Also the working relations between the SBR staff and the survey departments may be cumbered.

A coherent description of an integrated system for business statistics production which contains a statistical business register cannot treat the latter as exogenous ("given"). To overcome the issues mentioned, decisive questions are:

- What data sources feed a statistical register?

- What mechanisms harmonise the statistical register with the needs of its users?

- Who is contributing to the maintenance of a statistical register and by what motivation?

- How can the efficient maintenance of a statistical register be organised?

Ways to deal with these issues have been presented to the Wiesbaden Group in 2012. At that time, the technical infrastructure did not assist their realisation. Meanwhile a new data base system has been put into force in mid-2014.

4. The new data base for the German Statistical Business Register in 2014

In the case of the German SBR it has proven essential not only to improve the quality of the characteristics stored in the SBR but also to enlarge the variety of this data, in order to strengthen the motivation of both the register and the survey staff to maintain and use the SBR. As each survey department deals with quite different aspects the demands for the contents of the SBR tend to be accordingly heterogeneous. For one survey department it may be important to have additional indicators of size (such as number of beds of accommodations for tourism statistics), for other survey departments information about secondary activities or about special aspects of the production (such as certain kinds of pollution or of waste for environmental statistics) are of special relevance. The new IT solution provided in mid-2014 allowed for such enlargements of the content of the SBR according to the needs of the survey departments.

The development of the "New SBR" in Germany put emphasis on enlarging the possibilities to store subject matter information in the SBR in several ways:

- Metadata about the different surveys

- Clear picture of the sample populations

- More characteristics of the survey participants

- Time series/historisation of the survey related SBR variables stores within the data base

The new data base architecture consists of a core SBR section where all basic information of the legal and statistical units in the SBR is stored. A separate section contains all information that is survey related[1]. The two sections are connected of course so that it is possible to access all survey related information (stored in the "Survey SBR") of a register unit and combine it with its basic information (stored in the "Core SBR").

Figure 2
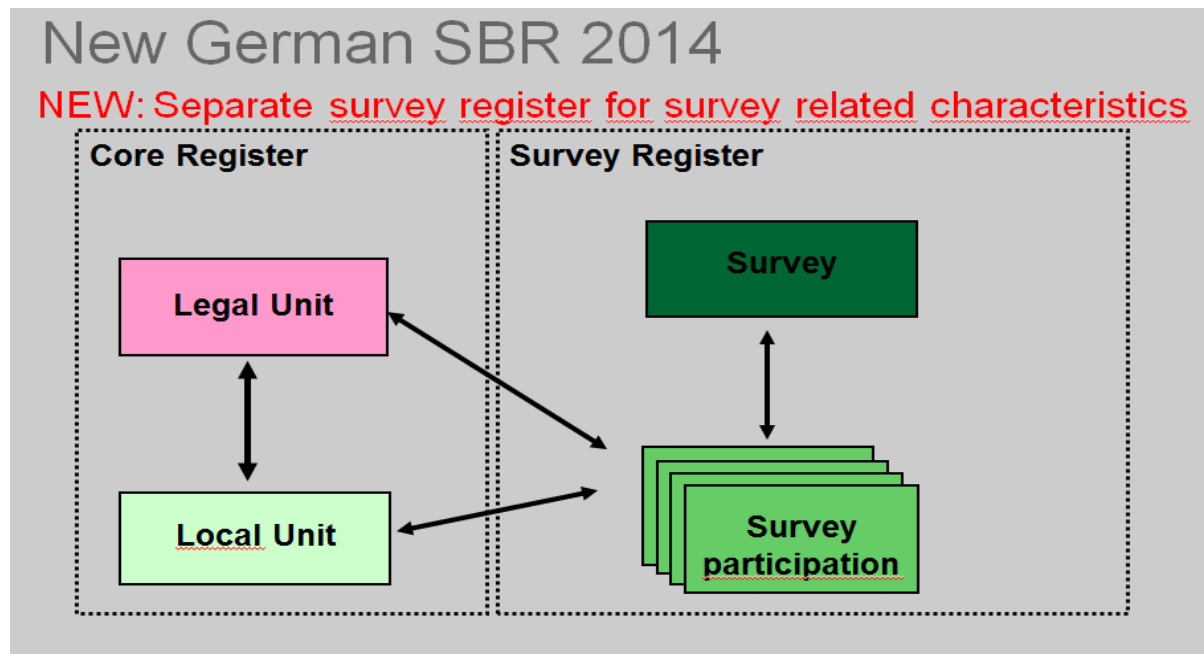
The Core SBR contains all basic data on each register unit – much of this is derived from administrative sources. Core SBR units are legal and statistical units. Examples for their characteristics are the name, legal form, size by turnover and employment, official address. Also demographic information and the links between register units are stored in the Core SBR.

The Survey SBR of the New SBR contains different kinds of information that are survey related – and often survey specific:

- Basic metadata of the different surveys themselves such as a code/identifier of the survey, the name of the survey, its centralised or decentralised way of conduction, information on burden management and some more are stored. The old SBR system contained basically the survey identifiers and related them directly to register units in the core SBR section. The name of the

---

[1] A third section of the SBR, which is not dealt with in this paper is the "Admin section" of the SBR, where all units from administrative sources are stored for purpose of SBR maintenance.

survey and all other meta-information about the survey had to be looked up in files outside the SBR or requested from the survey departments in charge of the survey. This made it difficult for users who wanted to get an overview of survey participation to assess the surveys that where indicated for a register unit.

- In the New SBR all survey participations of every register unit are dated by their reference year and stored over time in the IT data base, so that an overview of the past, present and planned future participations is at hand at any time. The old SBR system allowed storing only the survey participation of one reference year for each survey in the data base. The new system makes sure that no ambiguity about the reference year as well as the year of the actual conduction of the surveys can occur. This allows for an assessment of the overall response burden of the register units as well as for burden management.

- In the New SBR the address and contact information for each survey participation is stored. In the old SBR system only one address and contact information was available which had to be used by all affected survey departments– which proved insufficient and caused rivalries between the departments, about who was allowed store "their" contact information in the SBR. The disallowed departments in these cases stored "their" addresses and contact information in separate files. When new information from a survey corrected an older one, this was not reported sufficiently to the SBR department, even when the official address of the register unit was affected and should have been corrected.

- In the New SBR it is possible to store some predefined survey variables from every survey in which a unit participated. Basically these refer to turnover and employment. In the old BR system only one field for this survey data was available. Conflicting information was blocked from the SBR by priority rules. The new enlargement allows to detect inconsistencies between administrative and survey information as well as inconsistencies between the survey information coming from different surveys. This should encourage cross checking and cooperation between the survey departments.

- In the New SBR it is possible to store for every survey participation a number of variables which are not predefined. These "open characteristics" are for free use of the survey departments. By these variables some information in the SBR becomes explicitly survey specific. It is up to the survey departments to decide whether they want to make use of this possibility and which variables it should be. By this element the inequality of the old SBR system has been overcome which rejected many special demands of the survey departments in the past. It requires of course clear description and instruction for adequate use of these "open characteristics".
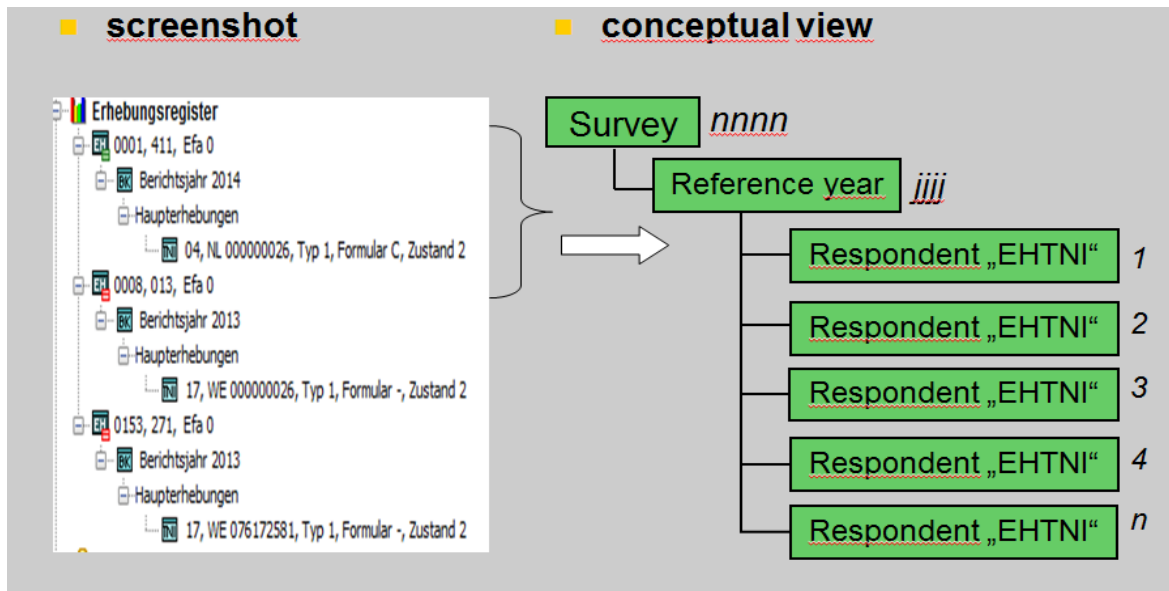
Figure 3

Figure 4

5. The enlargement of the new data base of the German Business Register in 2015

End of year 2015 an additional layer has been introduced in the data architecture of the German SBR. This layer is situated in the Survey SBR. Now survey departments have the possibility not only to store detailed information about their survey populations, but to introduce and maintain survey specific subunits of the units they find in the core SBR.

- The sub-units are related to a survey participation of a register unit in the Core SBR. As register units are stored in the Core SBR and survey participations are stored in the Survey SBR there is a clear attribution of maintenance rights for the two kinds of units: Core SBR register units are maintained by register staff, survey subunits are generated and maintained by the survey departments.

- Core SBR register units are defined by the EU regulation 696/93 on statistical units and they can be used by several surveys. Survey subunits are defined by the survey departments and they are survey specific. The way they are defined, generated, maintained and used is determined by the department which conducts a specific survey.

Figure 5
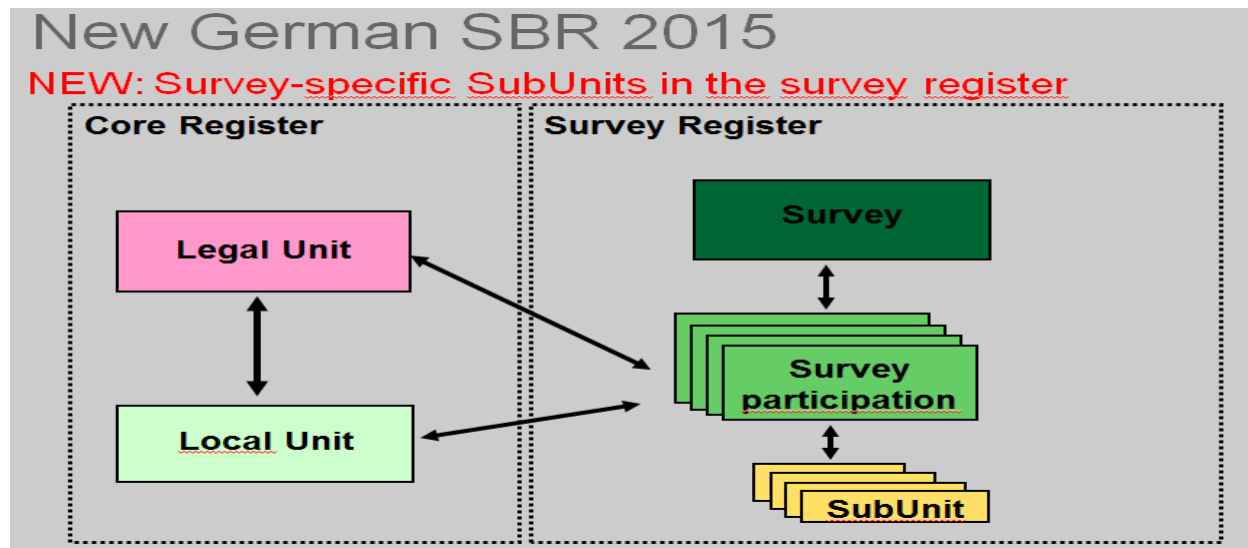
Survey departments can employ the survey sub-units for example to sub-divide a register unit in parts which do not fulfill the requirements of the register units of the Core SBR. For example they can sub-divide an enterprise in the tourism business, which provides rooms, meals and runs also a camping site into the sub-units "hotel" "restaurant" and "camping site", although these three might neither be separate enterprises or legal or local units and therefore cannot be represented as separate register units in the Core SBR. Another example may be an enterprise which has to report to pollution statistics separately for its dumping ground or its wastewater treatment plant. So these aspects of the survey can be described in the Survey Register by use of sub units.

This new feature of the SBR allows not only for storage of richer information, but allows also to better distinguish between reporting units, observation units and statistical units. The sub units in the Survey SBR commonly have the role of observation units and/or statistical units.

Figure 5
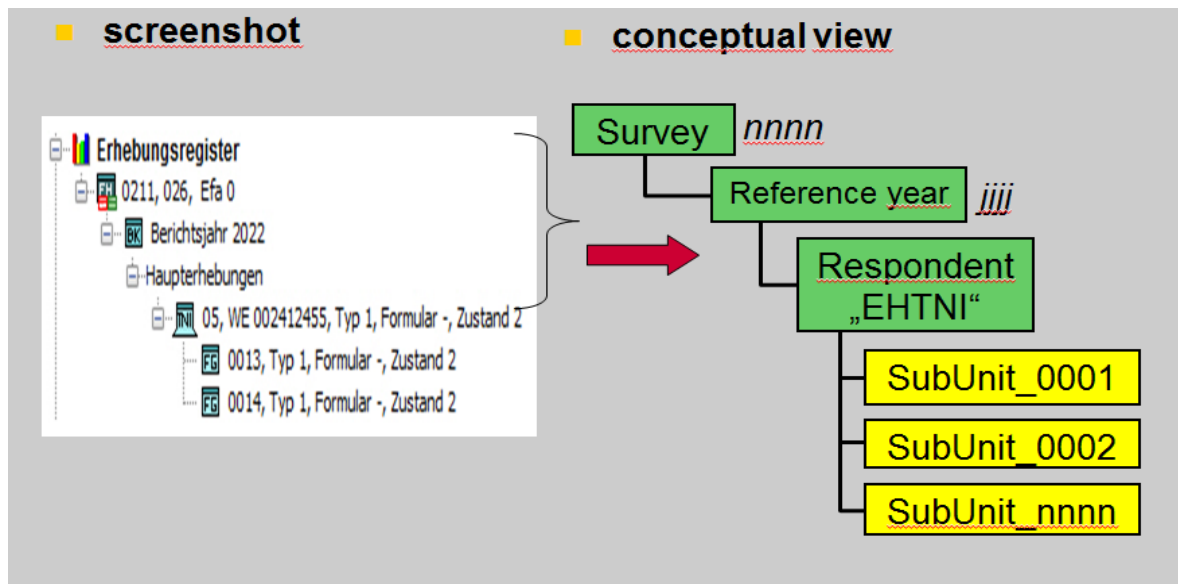
9. References

Sturm, R. (2007): Business Register as a tool for multi-source-analyses; paper presented at the 20th International Roundtable on Business Survey Frames; session 4; Wiesbaden, 23. 0ctober 2007

Sturm, R. (2008): The role of the Business Register in modern Business Statistics; STATISTIKA 6/2008, 471-482

Sturm, R. (2009): Konzept einer registerbasierten Wirtschaftsstatistik, in: Wirtschafts- und Sozialstatistisches Archiv, 4, 115-139

Sturm, R. (2010): Improving Consistency of Register and Survey Data, presented at the 2010 Business Data Collection Methodology workshop BDCM 2010 Wiesbaden, Germany, 28-30 April 2010

Sturm, R., Walsdorfer, K. (2010): Upgrade and re-design of the German business register data bank, presented at the 22nd Meeting of the Wiesbaden Group on Business Registers, Tallinn 25.-28. September 2010

Sturm, R. (2012): Organising data and organizing people for Business Register maintenance, presented at the 23rd Meeting of the Wiesbaden Group on Business Registers, 17.-20. September 2012, Washington D.C.

Sturm, R. (2014) Revised definitions for statistical units – methodology, application and user needs; paper presented at the Conference of European Statistical Stakeholders, Roma, November 24.-25. 2014