

Dated - Undated - Outdated: The Issue of Time Stamps and Reference Dates in Business Register Data”

Roland Sturm¹

Roland Sturm, Federal Statistical Office Germany, Gustav Stresemann Ring 11, DE-65189 Wiesbaden, Germany

roland.sturm@destatis.de

Abstract

A crucial feature of Business Register data is the attribution of time references to the register information. For appropriate use of register data it is crucial to distinguish clearly between the “technical time references” (the date/period of time an information enters the register or is edited in the register or made available by the register) and the “validity time reference” (the date/period of time for which the information is valid in the “real world”). Without “validity time reference” one may compare apples with oranges. The presentation will illustrate the complexity of the issue by some examples and show an approach to “do things properly” – which unfortunately leads to issues of complexity again.

Key Words: Business Register, time references, validity

1. The importance of reference dates

The world is changing. This statement is as much obvious – even trivial – as it is essential – even crucial – for the statistician who draws a picture of the real world by producing figures to describe it. Without reliable reference which state of the real world is pictured, the picture becomes unable to be interpreted, the figures unable to be used properly by decision makers.

Illustration 1: Business Register copy for analytical purposes

Business Registers in the Member States of the European Union have to fulfill the standards which are laid down in the EU Business Register Regulation. In Article 8 this regulation requires that

“Member States shall make annually a copy that reflects the state of the registers at the end of the year and keep that copy for at least 30 years for the purpose of analysis” (Article 8 (4) Regulation (EC) No 177/2008).

Eurostat explained to the Member States in the Working Group on Business Registers and Statistical Units that this legal wording “requests a copy reflecting the state of the register at the end of the year. The paragraph does NOT define:

- the state of the register with regards to a certain reference year,
- the day the snapshot shall exactly be taken or
- what day at the 'end of the year' the data refers to.

¹ Roland Sturm, Federal Statistical Office Germany, Gustav Stresemann Ring 11, DE-65189 Wiesbaden, Germany

roland.sturm@destatis.de

Assuming that one day in the second half of December is equal to 'end of the year', the results of the copy according to annex 1 (green arrow) would not dramatically differentiate for a respective MS.

Concerning the 'state of the register', from current legal point of view, two interpretations can be derived:

- MS may provide annually the snapshot (at end of the year) of the previous year.
- MS may provide annually the snapshot (at end of the year) of any other than the previous year.

The mixture of the legal opportunities result in the current existing situation, that – following the three examples in annex 1 – the snapshot has different reference dates and periods between countries BUT also within a country between different variables.

For the purpose of compliance towards Regulation (EC) No 177/2008 it is not relevant when the snapshot is taken as long as the snapshot includes data from a day 'end of the year' (green arrow in the figure of annex 1). From legal point of view it is not relevant if the snapshot is taken at the end of the previous year or one of the years before”. (cited from: DG Eurostat Unit G3 “SBR data quality programme Quality improvement – Q1” Date: 01/02/2016 Doc. Version: v1.0.)

Figure 1

2014 frame at different points in time for the purpose of Article 8

Examples for the updating of enterprise information	...	2014	2015					
		Dec	Jan	Feb	Mar	Apr	May	Jun
Member State 1	Legal status	Active	Active	Active	Active	Active	Active	...
	Reference date	30 11 2014	31 12 2014	31 1 2015	28 2 2015	31 3 2015	30 4 2015	...
	Turnover	100 €	100 €	100 €	80 €	80 €	80 €	...
	Reference date	2013	2013	2013	2014	2014	2014	...
	Employees	10	11	11	10	10	10	...
	Reference date	30 11 2014	31 12 2014	31 1 2015	28 2 2015	31 3 2015	30 4 2015	...
...								
Member State 2	Legal status	Active	Active	Active	Active	Active	Active	...
	Reference date	31 12 2013	31 12 2013	31 12 2013	31 12 2014	31 12 2014	31 12 2014	...
	Turnover	100 €	100 €	100 €	100 €	100 €	100 €	...
	Reference date	2013	2013	2013	2013	2013	2013	...
	Employees	15	15	15	12	12	12	...
	Reference date	2013 average	2013 average	2013 average	2014 average	2014 average	2014 average	...
...								
Member State 3	Legal status	Active	Active	Active	Active	Active	Active	...
	Reference date	31 12 2012	31 12 2012	31 12 2012	31 12 2013	31 12 2013	31 12 2013	...
	Turnover	150 €	150 €	150 €	150 €	100 €	100 €	...
	Reference date	2012	2012	2012	2012	2013	2013	...
	Employees	18	18	18	15	15	15	...
	Reference date	2012 average	2012 average	2012 average	2013 average	2013 average	2013 average	...



Legend	
Content in red	Information of a characteristic is updated
Content highlighted green	Info. of characteristics for Art.8 purposes
Content highlighted yellow	2014 info. of a characteristic is final

(figure cited from: DG Eurostat Unit G3 “SBR data quality programme Quality improvement – Q1” Date: 01/02/2016 Doc. Version: v1.0.)

What does this illustration show?

- The interpretation by Eurostat of the legal situation proposes that the annual copy could be a snapshot as illustrated in figure 1.
- The interpretation proposes that the “state of the registers at the end of the year” should be understood as “the end of a calendar year” and not as “the end of a reference year” (correctly, nothing else is a snapshot).
- The interpretation suggests that a snapshot could be used for purpose of analysis! But what analysis could base on such a mixture of information? Certainly not economic analysis. Probably analysis about the state of the art of the Business Registers in the Member States. Interesting as that would certainly be – but why keeping this documented for thirty years?

Eurostat intends to overcome this unsatisfactory situation – which surely is to be welcomed and supported – by the proposal that “Member States should make annually a copy – in April of a year (t) at latest – that reflects the state of the registers at the end of the previous year (t-1) and keep that copy for at least 30 years for the purpose of analysis.”

But still: It would result in a register copy that snapshots the “state of the registers” at a certain point of calendar time and take this as base information for analysis. This interpretation is not convincing.

Illustration 2: Business Register data in combination

The EuroGroupsRegister contains information on all multinational enterprise groups which are operating, at least partially, in the European Union. The production cycle of this register involves the statistical offices and the central banks of the 28 Member States of the European Union, two commercial data providers and Eurostat, where the data bank is hosted and which coordinates all the processes. The production cycle aims to produce annual frames for Foreign Affiliates Statistics (FATS) in the European Union and shall be extended to further uses.

The annual frames thereby combine data of (potentially) 58 sources. The data consists of three types of units which are interrelated:

- Legal Units (LeU) belonging to Multinational Enterprise Groups
- Enterprises (ENT) which comprise one or more Legal Units in one Member State of the European Union
- Multinational Enterprise Groups (so called Global Enterprise Groups GEG) which comprise more than one LeU in more than one Member States of the European Union.
-

The set of LeUs and the ENTs should be consistent with the Legal Units in the Statistical Business Registers in the statistical offices and the central banks in the member states. The GEGs are built from the national contributions and the contributions of the commercial data providers.

In the real world economy the set of legal units which belong to a GEG are subject to permanent change. LeUs are acquired or sold or created or liquidated according to the

development of the overall strategy of the MEGs. Information about “who belongs to whom” can be drawn from administrative registers and registers of the central banks, from the Annual Reports of the MEGs and in some Member States also from surveys of the statistical offices.

From this description it should be obvious that the production process of the EGR must be a sophisticated one to select the best information from all sources and at the same time a robust one to work efficiently – ideally it should be a clever combination of both aspects.

From the description it should become also clear that all sophistication and all robustness of the production of the EGR frames can only be successful if the data combined for the frame for one reference year really refers to that reference year. This again will depend on the production processes of the mentioned 58 sources and their ability to make sure that the reference dates of their data are correct. Whether this assumption holds – having in mind the illustration 1 and knowing that the annual EGR cycle starts five months after the reference year – may be assessed by all the 58 institutions participating in the EGR process.

2. Time references in the context of statistical business register maintenance

Statistical Business registers are fed by different sources. Often they combine information from one or more administrative sources (such as tax data, business registration data, social security data) and information from surveys. The data from these sources should contain information on which state of reality it describes (“validity time reference”, i.e. the date/period of time for which the information is valid in the “real world”). Maybe a different story is the “technical time reference” of the data, i.e. the date/period of time an information enters the register or is edited in the register or made available by the register. As these two time references usually do not coincide it is important to process the data appropriately in the maintenance of the statistical business register and – as users of the register information – to know about the maintenance procedures.

2.1 Validity time reference

This should be the easy one. The employment of a factory in December 2014 or the turnover of a business in year 2014 should be an unquestionable piece of data. Unfortunately it is not:

Are we talking about the employment on 31st of December 2014 or the average in the whole month?

Is the business aware about its turnover in the year 2014 if we ask this information on 31st of December 2014, or only three months later when most of the invoices should have brought money to the bank account and all customer complaints and all cancellations have been taken into account?

While the first issue can be solved by properly defining the data, the second issue depends on sensible consideration of the production process of the data.

2.2 Technical time reference

Technical time references may derive from many circumstances of the handling of data:

- When is a survey conducted?
- What is the time lag between tax declaration and its processing?
- How long does it take to announce and revise a turnover tax statement for a certain month till it reaches good coverage resp. becomes reliable?

When the German statistical business register altered its processing of tax data on turnover and social security data on employment from annual data to monthly data the latter questions had been examined. This led to the decision to wait six months after the reference month until processing the data. Tests had proofed the sensible time lags between first knowledge of information and sensible point of time to process it for register maintenance:

- Six calendar months after the reference month high coverage of turnover data is available (e.g. in September 2015 good coverage for turnover of March 2015 is available from tax administration records). For reasons of manageability it has been decided to process this data at one point of time for a certain reference month instead of several points of time according to availability.

Table 1: availability of turnover data (based on real data)

Availability of turnover from tax authorities (available turnover in % of total turnover)							
time lag	t+30	t +60	t+90	t+120	t+150	t+180	
Available turnover (%)	30	90	95	98	99	100	

- Although good coverage is available earlier for employment data, not earlier than six calendar months after the reference month high reliability of employment data is available from social security records (e.g.in September 2015 the figure which is supplied by the social security for March 2015 is more reliable than it had been in May 2015 when a first record on March 2015 becomes available).

Table 2: stability of employment data (based on real data)

Stability of employment figures from social security authorities (employees on local unit level)			
unit	t+60	t+90	t+180
1	1018	1010	1009
2	1007	1009	1008
3	991	995	996

Whereas economic data like turnover or employment need some time to become fit for use, the information on names and addresses can be processed without time lag between notice and processing.

The German business register data base contains for each variable both kinds of time reference: validity time reference and technical/processing time reference. Thereby the users have a clear picture about the content of the register. It is sometimes challenging to make proper use of this information. This is assisted by the technical search assistant that is used to select data from the business register: The technical assistant requests that both time references are filled in in order to start to select the data.

3. Purposes of a statistical business register and aspects of quality

Registers of different kind traditionally serve statisticians as means to generate survey frames. From this basic function statistical registers have been developed which offer assistance for different steps of the statistical production process. Modern business statistics in many statistical systems rely on a statistical business register as an infrastructural, conceptual and methodological backbone. This puts the statistical business register in a most responsible position:

- It shall contain all relevant statistical units and all relations between these units.
- It shall serve as frame for statistical surveys (for coherent sampling and extrapolation), therefor contain basic economic information for all its units (economic activity, size by turnover and employment).
- It shall be a tool to acquire, combine and connect existing data, therefor contain the various administrative identifiers of all its units.
- It shall be an assisting instrument for the statisticians in every phase of the production process of business statistics, therefore keep track with changes in reality and at the same time allow to view different states of reality retrospectively.
- It may serve as source for statistics in its own right.

As stated, the growing variety of purposes of the statistical business register brings about a variety of requirements which may not be harmonized. From the different and sometimes competing user needs it should be derived how a register is built up and maintained.

Figure 2 illustrates the requirements a survey department may have and the register information which is available for the assistance of the production process of business statistics. Traditionally the primary function of the statistical business register has been to provide the population of statistical units from which frames and samples for surveys can be drawn. If the survey statisticians make use of the register information available at different stages of conduction of the survey (sampling, sending out survey forms, extrapolation of sample results) as in the chosen example² they will face different stages of the maintenance work of the business register. The following description is based on the scenario that a statistical survey for reference year 2016 is to be conducted:

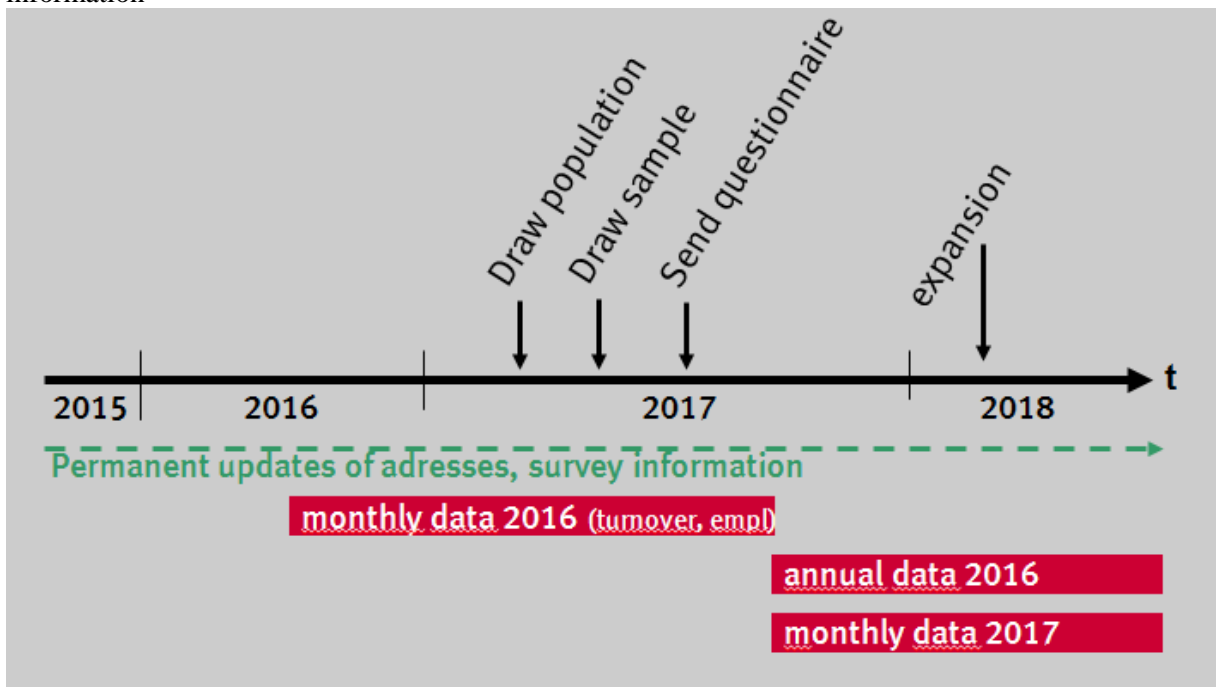
- For drawing a sample in March 2017 for reference year 2016 the survey statisticians will find in the register final annual data about the population 2015 and monthly data for January to June 2016. Thereby the register contains the total business population as known in June 2016 as base for the

² This example is based on register processes with regard to the German statistical business register. The analysis should hold for other comparable statistical registers as well.

determination of reporting units. As far as these units have existed in 2015, the survey statisticians can also read annual data on turnover and employment – referring to reference year 2015.

- For the stratification of their sample the survey statisticians can now choose whether to base this on annual data referring to 2015 – as far as the units have been active in 2015 or on an extrapolation of annual data for 2016 by monthly data for the first six months of 2016.
- When the survey statisticians send out their questionnaires in July 2017 they can do this on the base of address information relating to May 2017, as this is the most up-to-date information available from administrative sources which feed the register at that point of time.
- When the survey questionnaires have been processed and are ready for extrapolation in February 2018 the register contains the complete high quality population for reference year 2016 as well as annual economic data for reference year 2016.

Figure 2: stages of survey conduction and availability of business register information



4. Conclusion

To be able to make adequate use of all information in the statistical business register as early as possible it requires knowledge about the validity time reference of the register data.

Validity time references should be as disaggregated as possible – ideally available separately for every variable instead of for units in the register as a whole or even the complete register.

Technical time references of the maintenance procedure of a statistical business register have to be chosen according to requirements for quality (e.g. coverage or reliability) and the manageability of processes.

Snapshots of a statistical business register mix different validity time references and are not proper for economic analysis.

Frozen copies of a statistical business register which contain consistent information about a validity time reference will become available only a certain – often considerably long – period after the end of this reference time.

References

Philipp, K., Sturm, R. 2007: The role of the Business Register in reforming German Business Statistics, contributed paper for the Seminar on Registers in Statistics – methodology and quality, 21-23 May 2007, Helsinki

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

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