



This module is part of the

Memobust Handbook

on Methodology of Modern Business Statistics

26 March 2014

Theme: Building and Maintaining Statistical Registers to Support Business Surveys

Contents

- General section 3
 - 1. Summary 3
 - 2. General description..... 3
 - 2.1 The statistical registers 3
 - 2.2 The European framework of business registers..... 4
 - 2.3 The sources of the register maintenance..... 5
 - 2.4 The attributes of the registers 7
 - 2.5 The register maintenance..... 9
 - 2.6 The connection of statistical registers 10
 - 3. Design issues 11
 - 4. Available software tools 11
 - 5. Decision tree of methods 11
 - 6. Glossary..... 11
 - 7. References 11
- Interconnections with other modules..... 13
- Administrative section..... 14

General section

1. Summary

For business statistical surveys, the *statistical business register* (BR) and its connecting *satellite registers* are used as the main base for their *survey frames*. The purpose of these statistical registers is to record and maintain the *statistical units* and their characteristics (identifier, contact, classification, size category, etc. information) for a scope as complete for the economy as possible. The BR has an important role in the European Statistical System because it gives the opportunity to harmonise and integrate European statistics. Usually administrative sources are used for the *register maintenance*, but statistical surveys and other inputs can also amend information coming from these sources. The *frequency of the register maintenance* has to be harmonised with the periodicity and *reference period* of the surveys built on them.

2. General description

2.1 The statistical registers

The main sources of the *survey frames* are the statistical registers, in the business statistics this is primarily the *statistical business register* (BR).

In general sense the *register* according to the Business Registers – Recommendation Manual (Eurostat, 2010) is *”a written and complete record containing regular entries of items and details on particular set of objects. Administrative registers come from administrative sources and become statistical registers after passing through statistical processing in order to make them fit for statistical purposes (production of register based statistics, frame creation, etc.)”*

Based on this definition, administrative and statistical registers can be differentiated (UN ECE, 2012). The administrative registers are registers containing information not primarily collected for statistical purposes (e.g., administering taxes, benefits, realties or monitoring populations). It can be very useful, if the statistics can use them directly as input for the statistical surveys and data processing.

Administrative registers are, however, usually not suitable for the statistical surveys without additional control and correction. For statistical purposes, their populations might not be complete and additional variables are sometimes needed besides already available attributes in them. In such cases the administrative register cannot be used as a single source of the statistical register, only in combination with other sources.

A register can be used for various purposes:

- The primary role of the register is to record and maintain units that can become potential units of the survey populations. Consequently, the register serves as an input for the creation of the survey frame and statistical processes.
- Registers are also products of the statistical processes and provide a base for the analysis of the register population and the demography of the *register units*.
- Besides, they are tools for the coordination between statistical and administrative sources as they are the link in the integration of data coming from different sources.

Statistical registers are built for statistical purposes. They collect, store and maintain the units of a given population. Different criteria have to be satisfied by a register:

- It has to solve the regular update of the units.
- It has to describe the attributes for identification and accessibility of population units.
- It has to describe the attributes for supporting the surveying process of the population.
- It has to manage the linkage of the register unit with the units of other registers connected to it.
- It has to contain and maintain the current and historical statuses of the population and the causes, effects and sources of alterations in the population.
- Register data of population units have to be stored in a structured database.

If a register does not meet these criteria we can speak only about a registry or address list. Sometimes these are also suitable for the purposes of data collection, but it is quite rare. The aim is to develop the statistical register of all types of population units applied in the statistical surveys.

Example: The settlement register of the country records the settlements and their attributes. It follows the changes in the units and their attributes and gives possibility to indicate the given status of a settlement at a given date. In business surveys the settlement unit has analytical role, but the register is needed for the organisation, control of the surveys and the process and analysis of the collected and estimated data.

Regarding the topic of registers used in business statistics the registers can be classified into two groups:

- the first group describes the actors, units of the economy and their activity,
- the other group describes the geographical units.

Registers in the first group have significant role in the business surveys as they record the *statistical units* (see also “Statistical Registers and Frames - The Statistical Units and the Business Register”) either for the whole or a special part of the economy. The aim is that most part of these registers should be integrated into the BR. The harmonisation of the surveys both nationally and internationally necessitates the common interpretation of the unit types, the complete, unique and duplication-proof description of the units.

In the statistical data processing, the registers have an important role not only to describe the actors of the economy but also to record the *geographical and administrative units*. To fulfil this role, registers are usually designed to describe information on statistical units such as countries, settlements or addresses. These registers are connected to the BR and its satellite registers to control and locate the addresses and to group, analyse the relating statistical data by geographical units.

2.2 *The European framework of business registers*

The *Business Register Regulation* (EC) No 177/2008 of the European Parliament and of the Council establishes a common framework for these registers in Articles 1 and 3(1):

“Member States shall set up one or more harmonised registers for statistical purposes, as a tool for the preparation and coordination of surveys, as a source of information for the statistical analysis of the business population and its demography, for the use of administrative data, and for the identification and construction of statistical units.

The registers shall be compiled of:

- (a) all enterprises carrying on economic activities contributing to the gross domestic product (GDP), and their local units;*
- (b) the legal units of which those enterprises consist;*
- (c) truncated enterprise groups and multinational enterprise groups; and*
- (d) all-resident enterprise groups.”*

According to this Business Register Regulation, the mandatory *statistical unit* types of the register are the *enterprise group*, *enterprise* and *local unit*. The connection with the *legal units* is also crucial point of the register (EEC, 1993, UN DESA, 2007).

The incorporation of the other statistical units like *kind of activity unit* (KAU), *local kind of activity unit* (LKAU) is optional according to the possibilities of the countries. The remaining statistical units like the unit of the *homogeneous production* (UHP) and the *local unit of the homogeneous production* (local UHP) are analytical units, which have no direct counterpart in observable reality, therefore they are not part of the BR.

The regulation prescribes the *target population* of the BR, but some country-specific factors can cause coverage problems.

- The legal frameworks of the BR in the countries can be different, therefore, different obligations are in place to maintain the register.
- The available administrative registers that can be used as sources of the BR are also quite different in the countries. If there is no administrative source to maintain a given part of the population or a unit type of the register then it is very costly to provide a complete, up-to-date register with good quality.

In spite of the difficulties, the harmonisation of the statistics both at national and international level requires the establishment, maintenance and completeness of a BR.

Besides the national BRs, the EuroGroups Register (EGR) is creating a network of BRs used for statistical purposes in the Member States, focused on multinational enterprises. The EGR is foreseen to be the platform supporting the production of micro-based **statistics on globalisation** in Europe. The data will serve the needs of national statistical institutes and national central banks to compile statistics and will not be disseminated by Eurostat to the public.

2.3 *The sources of the register maintenance*

The *frequency of the register maintenance* is a crucial point in the timeliness of the register. The frequency of the register maintenance is the time interval of the register content alterations. Administrative registers are usually updated directly. In that case timeliness depends only on the time gap between the date of the real change and the registration. The update of the statistical registers may

be parallel with the administrative register sources, but they are usually maintained with certain periodicity. Sometimes the register can be maintained from different source with different periodicity. In that case the most frequently used source determines the frequency of the register maintenance. This periodicity of the maintenance (therefore the timeliness of the register units) is in close connection with the reference time of the surveys built upon the register. The aim is that the minimal periodicity of the register maintenance should be equal to the periodicities of the related surveys.

The register sources can be administrative registers, register surveys, statistical data transfers, feedbacks from survey results and other sources.

- The main sources of the statistical registers are the *administrative sources*. If there are administrative registers, registries available on the given topic, using these records is usually the most efficient solution considering time and cost restraints. This solution also puts fewer burdens on the respondents.

For the BR, the most important source is the VAT register that is available in all countries at taxation authorities. Besides the registers of the companies, the trade registration authorities, social security administrations and chambers of commerce can also be used. The used sources are different according to the peculiarities of the countries, particularly the legal system and the existing traditions. There are countries and authorities where certain administrative sources can be mobilised with difficulties only. In other countries the registration is operated by a so called “one-window” system where the unit is recorded in one authority and other authorities connected to the “one-window” and the statistical register can see and use directly the registration information.

- Another type of register source is the *register survey*. The primary purpose of these surveys is to collect information to update register information. They can be used to control the quality of the register, to get information on the activity/inactivity of the units or the changes of their certain attributes. For quality control, the register surveys are generally sample surveys.

Sometimes a certain part of the population is observed in full-scope by questionnaires, if there is no available administrative source to get information. The target population of these surveys can be, for example, the newly registered companies, large companies or subpopulation with a given activity or legal form.

Example: A register for the legal units is available from administrative sources, but it doesn't contain proper information about the local units. In that case a register survey can be applied for the legal units (or the legal units above given size categories) to collect data about the local units. The survey frame of the register survey is built on the register of the legal units. The questionnaire asks the legal units to enumerate their local units or the changes in their local units.

Full-scope (or census type) surveys for the whole population are very costly therefore they are conducted only if there is no administrative source for the topic at all (in special cases to get information about the new units of the register this type of surveys can be built on the address register as a survey frame).

- The third type of register sources is *the transfer of statistical data*. The register unit can have attributes updated from statistical sources. For example, in the BR such attributes are the

number of employees, the value of the turnover or the *principal activity* of the units. As only a small part of the BR units is observed by statistical surveys, survey data are not available to maintain these attributes for the whole register population. The attributes of the non-observed units should be maintained from administrative sources. Survey data can provide more up-to-date, accurate information than the administrative data for the observed units.

- For register quality, *feedbacks from the statistical surveys* can also be important sources. The data collection is built on the survey frame that is based on the register. During the data collection phase, accessing the respondents might reveal the error of the survey frame attributes, the changes of the address, name, the activity state, activity type or other attributes of the respondents. The same feedback can be gathered from questionnaires returned. The feedback information can be used in the survey frames of the next reference periods to improve their quality, but its influence to the sample and the estimation have to be considered (not to introduce a bias of the estimates based on the register).
- As an addition to the sources mentioned above, some *other sources* of information can also improve the completeness and punctuality of the register. Information from different communication channels, media about the birth and changes of register units (such as brochures, publications) are typical examples of this kind of sources. Information directly from the units in charge of reporting about the changes in their data can be another example. As this kind of information is not complete, this source has only a supplementary role in register maintenance.

The information coming from the different sources can have different priority and role in the update, in creation of the survey frame and later in the statistical process. These sources affect the completeness and timeliness of the register. For more information on how to control and measure the quality of the statistical registers see the module “Statistical Registers and Frames – Quality of Statistical Registers and Frames”.

2.4 *The attributes of the registers*

The register units have to be characterised with the attributes that support the functions of the register maintenance and usage. The main functions of the registers are:

- to collect and store information about the register population from one or more sources,
- to provide frame for collection and processing of statistical data, and
- to facilitate the analysis of the register population demography.

For these functions not only the units of the register but their attributes have to be updated regularly.

The attributes of the register can be grouped according their roles:

- *Identification characteristics* uniquely identify a given register unit (an occurrence of a given register unit type). Usually this is an identification number, identifier. The name and address of the register unit are also considered as identification characteristics, but they can only have secondary importance in the identification in the absence of an accepted identifier. Their main purpose is rather to give contact information to the units.

- *Contact characteristics* provide information to reach the unit for surveys. These attributes are usually the name, address, the boss of the unit, the contact person, their postcode, phone-, fax numbers and e-mail addresses.
- *Demographic characteristics* show the formation, cessation (birth, death), activity, inactivity, the type and date of changes, the current activity state and other changes of the register units.
- *Economic/stratification characteristics* support the classification of the units to range them into different strata for sampling, grossing and for the analyses. This type of attributes can be either classifications or size categories. In the BR the most important attributes are the *NACE*, the legal form of the unit, the geographical location (NUTS and LAU), the institutional sector of ESA, the categories of the number of employees and the value of turnover.
- *Link characteristics* facilitate the connection with other registers or sources. The link characteristic is usually an identifier of the unit that is used both in the given and the linked register. Such relation can be established with the tax system storing the VAT/tax number, or through the social security number with the social security register. Link characteristics can also be used to describe connection between different units in the same register. BRs usually contain information for legal units and enterprises therefore the legal unit can be characterised by the identifier of the enterprise it belongs to.
- *Control- and ownership characteristics* identify the units in the given or another register with control or ownership over the given unit of the register.
- *Maintenance characteristics* support register maintenance only, usually to identify the source of information and describe the date and effect of the register update.

The different types of statistical units (enterprise, legal unit, local unit, etc.) in the register (registers) are in close connection with each other, therefore the identification and maintenance system has to manage their coordinated linkage:

- by the application of common identifier in the connecting register units or
- in the case of different identifiers by the maintenance of the connection between the identifiers of the register units,
- the harmonisation of the frequency of the register maintenance as much as possible and
- the maintenance of the link and state of register units at the change of the other register unit.

As it was mentioned earlier the administrative and the statistical registers can be differentiated. In the attributes of the register differences can also be made in administrative and statistical attributes.

- There are attributes that come from administrative sources and cannot be modified as they are considered to be the official data of the given units. Such attribute can be, e.g., the name and seat address of the legal unit.
- There are attributes that come from statistical sources like the number of employees. These attributes can be modified if they are inaccurate.

- There are attributes where the default values come from administrative sources, however, their quality essentially determine the quality of the statistics. In that case the modified (checked and corrected) values of these attributes are statistical attributes.

The administrative attributes can change according to the frequency of the register maintenance from administrative sources. The statistical attributes should be updated once a year, because it has to be stable for the comparability of statistical data in the different periods of the year.

Example: information on the principal activity of the legal unit comes from administrative sources. It can be controlled by the algorithmic rule of the principal activity built on the available statistical and administrative data of the legal unit. If the administrative activity does not fit the composed principal activity in a longer period, it can be substituted with the statistically accepted code value.

2.5 *The register maintenance*

The task of the register is to update the units and their attributes with the changes. The main types of demographic events (Eurostat, 2010) are the birth, death, reactivation of the units, concentration (merger, take-over), de-concentration (break up, split-off), changing of ownership and grouping of units. Besides the register maintenance has to manage, update the changes in the attributes of the units

Sometimes the sources do not give sufficient information to define the cause of changes or to define the correspondence of the units (if for example there is no common identifier of the register and the source of maintenance) or to define the continuity of the units. In those cases the maintenance demands a more complex application to manage the changes and to define the cause of demographic events. (See the topic “Dynamics of the Business Population”).

The register has to store the actual and historical state of the register units and the cause of the changes. The type of the change and history management of the registers can be different.

- The simple form of register maintenance is when a new register instance (a new data set) is created every year that contains the living units of the total population. The actual instance shows the current state of the units, but the previous states of these units (and dead units) can be accessed from the earlier instances, if needed.
- The best way of change and history management is when a register shows all units with their attributes that either exist or ever existed. Besides the current state, all changes in the content, type, date and source are registered. From historical information the attributes of the units can be accessed for any earlier date. At this register type the maintenance of the attributes can be ranged into two groups: one is when the attribute values are followed historically (therefore historical values are available for previous states) and the other one is when the register does not save the previous values of the attribute (therefore only the current values are available).

Example: in the BR the name and seat address have very important roles in the identification of the unit, the principal activity and geographical location in the statistical processing. Therefore the history of these attributes is saved. Postal address of the unit, for example, has only contact role therefore saving the earlier states of this attribute is not necessary.

2.6 The connection of statistical registers

In business statistics the national BR plays the central role. It should cover all enterprises with economic activities to as large extent as possible. That provides the unified identification, characterisation, maintenance of the actors of the economy and guarantees the control and exclusion of any duplication in the units. The best solution is when the BR also maintains other types of statistical units (local units, kind of activity units and local kind of activity units) besides the enterprises.

There are cases, however, when separate statistical or administrative registers manage a certain set and activity type of the actors of the economy (e.g., shops, accommodations, farms, R&D activity, Intrastat/Extrastat foreign trade activity, etc.). In these cases the links of these registers to the BR should be built, that is they have to store the identifier of the connected BR unit.

These registers have role in the control of the BR and vice versa, to control the activity/inactivity of the units. These types of registers are called associated or *satellite registers*. According to the Business Register Recommendation Manual (Eurostat, 2010):

“Satellite registers are therefore defined as registers that are available to the national statistical system, contain information about businesses, and fulfil the following conditions:

- *They are not an integral part of the statistical business register as referred to in the business registers Regulation, but are capable of being linked to it.*
- *They are more limited in scope than the statistical business register, e.g. in terms of NACE, but within that scope they may have more extensive coverage of units and/or variables.*
- *They contain one or more variables that are not found in the statistical business register. Such variables are generally capable of being used for stratification purposes”*

Instead of maintaining separate registers, it would be a better solution to integrate these registers into the BR. If there is no way to do so, then the harmonisation of the register content and management of the units is an important issue. The harmonisation or integration of the register has similar requirements as the linkage of different types of register units

- the application of common identifier in the connecting registers or
- in the case of different identifiers the maintenance of the connection between the identifiers of the registers,
- the synchronisation of the register attributes,
- the harmonisation of the frequency of the register maintenance as far as possible,
- the maintenance of the link and state of register units at the change of the other register unit.

Example: In Hungary the retail trade register records the shops and their activities. The retail trade register is not direct part of the BR, it is maintained separately. About the shops it records additional information like the type, activity and the size of the shops that do not characterise other types of LKAUs in the BR. The retail trade register has different sources and update frequency compared to the BR and the change management is also different. Therefore it is handled as a separate register. The retail trade register can be linked to the BR through the VAT number of the trader legal unit. That enhances change tracking of the legal units, because their maintenance could be more frequent than it is in the case of the shops.

Besides the satellite registers, another important linkage option for the national BR is the EuroGroups Register (EGR). The national BR has to be able to connect to the EGR with the legal units that belong to multinational enterprise groups. Even though this kind of role and task of BRs is growing, this ability is currently limited because of the available sources to define the place of the units in the hierarchies of the legal units within the enterprise groups.

3. Design issues

4. Available software tools

5. Decision tree of methods

6. Glossary

For definitions of terms used in this module, please refer to the separate “Glossary” provided as part of the handbook.

7. References

Council of the European Communities (1993), Council Regulation (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community. *Official Journal of the European Union*, L 076, 30/03/1993, 0001–0011.

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31993R0696:EN:HTML>

European Parliament and the Council of the European Union (2008), Regulation (EC) No 177/2008 of the European Parliament and of the Council of 20 February 2008 establishing a common framework for business registers for statistical purposes and repealing Council Regulation (EEC) No 2186/93. *Official Journal of the European Union*, L 61, 05/03/2008, 0006–0016.

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:061:0006:01:EN:HTML>

Eurostat (2010), *Business Registers - Recommendations Manual, 2010 edition*. Eurostat Methodologies and Working papers. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-32-10-216/EN/KS-32-10-216-EN.PDF

UN DESA (2007), *Statistical Units*. United Nations, New York.

<http://unstats.un.org/unsd/isdts/docs/StatisticalUnits.pdf>

UN/ECE, United Nations Economic Commission for Europe (2012), *Using Administrative and Secondary Sources for Official Statistics. A Handbook of Principles and Practices*.

http://www.unece.org/fileadmin/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf

Interconnections with other modules

8. Related themes described in other modules

1. User Needs – Specification of User Needs for Business Statistics
2. Overall Design – Overall Design
3. Repeated Surveys – Repeated Surveys
4. Statistical Registers and Frames – Main Module
5. Statistical Registers and Frames – The Populations, Frames, and Units of Business Surveys
6. Statistical Registers and Frames – Survey Frames for Business Surveys
7. Statistical Registers and Frames – The Design of Statistical Registers and Survey Frames
8. Statistical Registers and Frames – The Statistical Units and the Business Register
9. Statistical Registers and Frames – Quality of Statistical Registers and Frames
10. Dynamics of the Business Population – Business Demography
11. Sample Selection – Main Module
12. Data Collection – Main Module
13. Response – Response Process
14. Micro-Fusion – Data Fusion at Micro Level

9. Methods explicitly referred to in this module

- 1.

10. Mathematical techniques explicitly referred to in this module

- 1.

11. GSBPM phases explicitly referred to in this module

1. All phases for register maintenance

12. Tools explicitly referred to in this module

- 1.

13. Process steps explicitly referred to in this module

1. All processes for the register maintenance

Administrative section

14. Module code

Statistical Registers and Frames-T-Building and Maintaining

15. Version history

Version	Date	Description of changes	Author	Institute
0.2	29-05-2012	first version as a separate module	Ildikó Györki	HCSO
0.3	16-07-2013	revised version after EB review	Ildikó Györki	HCSO
0.4	17-11-2013	revised version after EB second review	Ildikó Györki	HCSO
0.4.1	11-12-2013	preliminary release		
0.5	25-01-2014	revised version after EB third review	Ildikó Györki	HCSO
1.0	26-03-2014	final version within the Memobust project		

16. Template version and print date

Template version used	1.0 p 4 d.d. 22-11-2012
Print date	21-3-2014 17:36