REPUBLIC OF SERBIA
STATISTICAL OFFICE OF THE REPUBLIC OF SERBIA

WORKING DOCUMENT

ISSN 1820-0141

SAMPLE COORDINATION OF STATISTICAL BUSINESS SURVEYS

www.stat.gov.rs

Year L Belgrade, 2014 N° 89

REPUBLIC OF SERBIA
STATISTICAL OFFICE OF THE REPUBLIC OF SERBIA

WORKING PAPER

ISSN 1820 – 0141

SAMPLE COORDINATION OF STATISTICAL BUSINESS SURVEYS

Year L	Belgrade, November 2014	Number 89

Publisher:
Statistical Office of the Republic of Serbia
Belgrade, 5, Milana Rakića St

For publisher:
Director
Professor Dragan Vukmirović, PhD

Foreword

The role of sample coordination of statistical business surveys is to ensure control of the overlap of samples in order to enable precise estimate of change, improve comparability of data from different statistical surveys, increase the efficiency of data production and at the same time reduce the burden on respondents.

The basis of the methodology on sample coordination of business surveys at the Statistical Office of the Republic of Serbia was set in 2013. Since then all surveys on random samples of enterprises are included in the system of coordination.

Implemented system of coordinating is based on the Swedish system SAMU for coordination of frame populations and samples from the Business Register. Experts Annika Lindblom and Stefan Berg gave their assistance and advice in setting the methodology within the Activity on Survey Methodology of the Project Partnership in Statistics funded by SIDA (Swedish International Development Cooperation Agency).

Finally, this document was prepared in the Component 19 that deals with business register of the Project MB_IPA 2011 IPA 2011 (Multi-beneficiary statistical cooperation program).

Presented methodology is of interest to those dealing with business surveys, statistical registers and sample survey methodology. For the employees of the Statistical Office of the Republic of Serbia it can be a starting point for further development of the system of coordination of samples and frame populations of all business surveys, not only of those on random samples.

Working paper *Sample Coordination of Business Surveys* is also presented in electronic form on the internet page www.stat.gov.rs.

Belgrade, 2014

Director
Prof. Dragan Vukmirovic, PhD

Acronyms

PRN	Permanent random number
SAMU	System for coordination of frame populations and samples from the Business Register at Statistics Sweden
SBR	Statistical business register
SORS	Statistical Office of the Republic of Serbia
SRSWOR	Simple random sample without replacement

Contents

1. Introduction

Many business surveys are conducted in official statistics with the aim to collect data that can be exploited in estimation of the level and the dynamics of the economy of a country.

Business surveys are usually based on samples that are selected from frames – lists of units which define a population to be sampled. Frames are most often constructed from the Statistical business register (SBR), but frames can be constructed from administrative registers and other lists or areas.

There are some desirable features of business surveys. For example, it is preferable that different surveys have common concepts, frames and sample designs. These characteristics enable production of comparable statistics across surveys for the same definition of units (enterprise, local kind of activity units, etc.) that is needed for national accounts. Also, efficiency of production of statistical data could be increased, as it is possible to use the single procedures for frame updating, contact information and editing of survey data.

Business surveys are numerous and most often repeated in time. They should be such that the burden on sample units is limited. Limited burden should be expressed in the number of surveys a unit is selected for, as well as in the amount of time a unit is in repeated survey.

On the other hand, the samples of repeated surveys should overlap for high precision of change estimates, minimized disruption of estimates in sample redesign and to avoid excessive costs of data collection.

To make a balance between these desirable opposed features of business surveys, official agencies of statistics use coordinated rotating sampling for business surveys.

A method of survey coordination developed at the Statistical Office of the Republic of Serbia (SORS) is based on the system for coordination of frame populations and samples from the Business Register at Statistics Sweden (SAMU).

Up to now, coordination at SORS has been implemented for business surveys of statistical enterprises that are based on random samples. The following has been achieved:

- Overlapping of samples of different surveys and repeated surveys is controlled.
- The burden for small enterprises that belong to strata with small sampling fractions (ratio of number of units in the sample and the total number of units in the stratum) is reduced.
- Comparability of statistics across surveys is increased.

After this introduction, short reviews are presented in parts 2 – 5 of the sequential random sampling, the JALES method for positive and negative sample coordination, blocks of surveys and sample rotation. Parts 6 – 7 are devoted to characteristics of sample coordination at SORS. Concluding remarks and suggestions for further work on coordination and reduction of statistical burden of enterprises are presented in part 8.

2. Sequential random sampling

Simple random sample without replacement (SRSWOR) is equivalent to sequential simple random sample Ohlsson (1992). For this reason a SRSWOR of size n from the population of size N can be selected by a following sequential scheme:

- To each unit k form SBR associate a random number u_k, $\mathrm{k}=1,\ldots,\mathrm{N}_{\mathrm{SBR}}$ from uniform distribution on the interval $[0,1]$;
- Order the frame population (subset of SBR) in ascending random number sequence;
- Select the first n units.

In the similar way a stratified simple random sample can be selected, in which case a sample of size n_h is selected from stratum h of size N_h using the described scheme.

Associated random numbers to SBR units can also be used for sequential selection of a Pareto πps sample. Pareto πps is a specific probability proportional to size sample without replacement. For more details see Appendix 1.

Sequential SRSWOR random sample of size n can be selected from any fixed point a in the interval $[0,1]$, by selecting n units to the right from a. If there are not enough units to the right of point a selection is continued to the right from point 0. By symmetry, selection can be conducted to the left of the point a.

3. JALES method for positive and negative sample coordination

Sample coordination or control of size of sample overlaps can be achieved by JALES technique that was developed at Statistics Sweden in the early 1970s. This method for sample selection is based on random numbers, associated with units from SBR, that are permanent (PRN), apart from being independent, unique and uniformly distributed on the interval $[0,1]$. With the proper choice of starting points, JALES technique will ensure the following:

- A sample selected from the updated SBR, with the same starting point and in the same direction as the previous sample of a survey, will have large overlap with the previous one – positive coordination.

  The size of overlap depends upon the dynamics of the population considered (births and deaths of economical units), changes in sample design (if they are introduced).
- Samples across surveys that are selected from the same starting point in the same direction will have large overlap – positive coordination.

  The size of overlap depends upon the differences in coverage of surveys and design of samples.
- Samples across survey that are selected from starting points well apart and in the same direction will have a small overlap – negative coordination.

The sampling fraction and probability of inclusion of any unit k in the sample, for the simple random sample without replacement, are equal:

$$f = \frac{n}{N}$$

where N is the number of units in the sampling frame and n the sample size. In sequential random sampling, random numbers associated with the sampling units because of uniform distribution, belong to the interval with length f (to the right/left of the starting point, depending on the direction of sample selection).

For surveys with different starting points for sample selections, a_1 and a_2, $a_1 < a_2$ and sampling fractions f_1 and f_2 maximized negative coordination will be achieved if:

$$f_1 < (a_2 - a_1) \text{ and } f_2 < (1 - a_2 + a_1)$$

This method of coordination applies to stratified simple random samples, as well. For all strata of a survey sampling frame, the same starting point is used and the sample is selected in the same direction in each stratum.

As for simple random samples, samples of different surveys that have different starting points and the same direction of sample selection will be negatively coordinated, and the ones that have the same starting point and the same direction of sample selection will be positively coordinated.

In case the surveys have the same coverage, regardless of the starting point, samples will overlap for strata with large sample fractions (strata with small number of units or census strata). This is most often the case for strata of medium and large businesses, due to skewed business population.

4. Blocks of surveys – positive and negative coordination

Successive samples of the same survey overlap maximally as they are usually selected from the same starting point and in the same direction. Owing to this, survey data collection costs are reduced. Also, this enables high precision of change estimates, particularly for short term statistics.

In order to get coordinated, surveys are grouped in blocks according to their starting point for sample selection. The grouping of surveys is usually done in a way that burden of sampling units is spread over frame units. Surveys that have different coverage of units (with respect to economic activity, size or territory) can be assigned to the same block without increasing the burden of enterprises. When sample of one survey is a sub-sample of another survey, both surveys have to be assigned to the same block.

5. Rotation

Successive samples of a survey with JALES method have large overlaps, with differences mainly caused by the dynamics of the universe of businesses. Units that were selected for the sample would be in the sample for a long time, while other units (not selected) would have a small chance of being selected. Therefore, it is necessary to introduce a system of rotation with JALES method of sample selection. At the same time, the rotation system should enable refreshing of the sample and certain percentage of overlap of the new with the old sample.

A method of rotation that is based on random partitioning of units from SBR according to their permanent random numbers into R rotation groups, approximately with the same number of units, ensures an equal rotation rate for samples of all surveys; thereby the coordination across surveys is maintained. By this method permanent random numbers of one rotation group are shifted annually, for a predetermined shift b. On the average, samples would be renewed each year with the rate of

$$1/R$$

and in R years completely. This is valid only for strata where there are enough units for this kind of change and those are the ones with the sampling fraction less than or equal b. Differences from expected rotation, under the same sampling design, are caused by dynamics of business entities (deaths, births and transformation of units).

For example, for strata with sampling fraction (inclusion probability of a unit in the sample) less than or equal 0.1 and the shift of a rotation group of 0.1:

- If the number of rotation groups is 5, it is expected that the sample changes by 20% each year and in 5 years completely;
- If the number of rotation groups is 3, it is expected that the sample changes by 33% each year and in 3 years completely.

The shifting of the permanent random numbers (PRN) is conducted in the following way:

PRN -> PRN+b (if PRN+b<1)

PRN -> PRN+ b-1 (if PRN+b>=1)

6. Characteristics of SORS business surveys based on random samples – Short review

Some characteristics of SORS business surveys that are based on random samples are the following:

- The population of interest is extracted from SBR data on statistical enterprises. So far, data of other types of units of SBR have not been used for construction of a population.
- In order to reduce the costs and non-response, the sampling frame is constructed from the population with cut-off. The smallest enterprises in domain of interest are cut off according to turnover, number of employees or some other measure of size. As small enterprises are numerous but with small contribution to the total turnover, for example, the number of units in the frame is much smaller than in the population, while at least 95% of the turnover remains covered.
- The stratification of the sampling frame is very detailed and in compliance with size, economic activity or some other characteristic that refers to the domain of interest. Defined strata are additionally stratified, in almost all surveys, according to whether they include the largest units (census strata) or not (strata from which random sample is selected). The census strata could contain up to 50% units of the entire sample.
- Permanent random numbers are used for selection of a sequential simple random sample (JALES technique). In several surveys Pareto πps sampling is used.
- Surveys are grouped into four blocks according to the starting points: 0; 0.25; 0.50 and 0.75.
- In all surveys, sample selection is to the right from the starting point.

  Sample coordination can be achieved with selection in both directions from the starting point, to the right and to the left. Selection of SORS samples is carried out in one direction, to the right of the starting point, for easier maintenance of sample coordination.
- Units of SBR are randomly partitioned into five rotation groups. Once a year PRN numbers of one rotation group are moved for a shift of 0.1 for all units that are not new in the current year.
- Rotation of enterprises started in 2013:
  - PRN numbers of the first rotation group were shifted for 0.1 in 2013;
  - PRN numbers of the second rotation group were shifted for 0.1 in 2014.

  This means that in five years, starting form 2013, all groups will have shifted permanent random numbers.
- The estimation procedure is carried out with weights corrected for non-response. In some surveys outlier units are detected and treated. Also, calibration to some known auxiliary domain totals is used only in several surveys.

7. Blocks of SORS business surveys based on random samples in 2013

The list of surveys on random samples by blocks, according to the starting point for sample selection, is given in Table 1. Table 2 contains a summarized presentation of the conducted surveys by periodicity and starting point for sample selection.

Table 1. Surveys based on random samples at SORS, by blocks, 2013

Row	Code	Name of the survey	Starting point	Block	Periodicity*	Reference period	Year of conducting survey
1	IKT	Annual survey on information technologies	0	1	A	2012	2013
2	IND01	Monthly industrial sample survey on turnover for small enterprises	0	1	M	M	2013
3	UG11p	Quarterly catering survey for entrepreneurs	0	1	Q	Q	2013
4	UG11	Quarterly catering survey for legal units	0	1	Q	Q	2013
5	INV01p	Annual survey on investments for entrepreneurs	0	1	A	2012	2013
6	OTGIS	WASTE generated in construction and services activities (OT-GIS) u 2012 – pilot	0.25	2	A	2012	2013
7	SBS03	Quarterly structural business survey	0.25	2	Q	Q	2013
8	IND02	Quarterly survey on turnover and new orders in industry	0.25	2	Q	Q	2013
9	TRG15p	Biannual trade survey for entrepreneurs – NEW	0.25	2	BA	BA	2013
10	ARAD1p	Biannual survey on number of employees in small enterprises	0.25	2	BA	BA	2013
11	INV01	Annual survey on investments in small enterprises	0.5	3	A	2012	2013
12	TRG10	Monthly retail trade survey	0.5	3	M	M	2013
13	TRG16	Quarterly retail trade survey	0.5	3	Q	Q	2013
14	TRG16m	Quarterly wholesale and retail trade of motor vehicles and motor-cycles	0.5	3	Q	Q	2013
15	TRG16kv	Quarterly wholesale trade survey	0.5	3	Q	Q	2013
16	INOV	Innovation survey (once in 2 years)	0.75	4	2A	2011-2012	2013
17	ITR	Labor cost survey (once in 4 years)	0.75	4	4A	2012	2013
18	BCind	Monthly business and consumer survey – industry	0.75	4	M	M	2013
19	BCusl	Monthly business and consumer survey – services	0.75	4	M	M	2013
20	BCgradj	Monthly business and consumer survey – construction	0.75	4	M	M	2013
21	BCtrg	Monthly business and consumer survey – trade	0.75	4	M	M	2013

* Periodicity of the survey**:** Monthly (M); Quarterly (Q); Biannual (BA); Annual (A); Once in 2 years (2A); Once in 4 years (4A)

Table 2. Number of surveys based on random samples by starting points and periodicity, SORS, 2013

Periodicity	Total	Starting point for sample selections			
		0	0.25	0.5	0.75
Total	**21**	**5**	**5**	**5**	**6**
Monthly (M)	6	1		1	4
Quarterly (Q)	7	2	2	3	
Biannual (BA)	2		2		
Annual (A)	4	2	1	1	
Once in 2 years (2A)	1				1
Once in 4 years (4A)	1				1

Tables 3a and 3b present the number of enterprises included in survey frames, by classes of employees and the number sampling frames which they belong to, the absolute numbers and the percentages. Similarly, Tables 4a and 4b present enterprises included in samples, by classes of employees and the number samples which they belong to. A detailed presentation of enterprises of business survey samples by size classes, number of survey frames and survey samples is given in Appendix 2.

In 2013, twenty one surveys based on random samples were conducted. Concerning the data on frame and sample units that were used for calculating results presented in tables 3a, 3b, 4a and 4b, some additional explanations are given below.

Biannual survey on number of employees for March was based on one frame and sample, while for September the frame was updated and a new sample was selected. For the tabulation, the frame contained all units that were in the union of the two frames. In the similar way the sample contained all units that were in at least one of the two samples. For the units that were in both frames, the number of employees and turnover were taken to be as in the September frame.

Estimation for the Quarterly structural business survey (SBS03) was based on enterprise data with one exception, Petroleum Industry of Serbia (NIS). The data for Petroleum Industry of Serbia was split into 11 parts according to NACE Rev. 2 classification of economic activity using the contribution of individual parts in the value added of Petroleum Industry of Serbia. In the frame and sample, Petroleum Industry of Serbia was presented by 11 parts. The Quarterly survey on turnover and new orders in industry (IND02) was based on the same frame, sample and questionnaire, but with data for Petroleum Industry of Serbia not split in parts. For the tabulation, SBS03 and IND02 were considered as one survey and Petroleum Industry of Serbia on the level of enterprise. In that way, results presented in tables 3a – 4b refer to 20 instead of 21 surveys of enterprises.

In all, 105 thousand enterprises from the SBR were in the large sampling frame – a list of units that were in at least one of the sampling frames for the conducted surveys on random samples in 2013. The SBR included 419 thousands of registered enterprises in February 2013, so that 25% of the SBR enterprises had a chance of being selected in one of the surveys that were based on random samples.

The distribution of enterprises of the large sampling frame by classes of number of employees was highly skewed. The most of the enterprises (73%) were in the class of 1 – 9 employees, 10% of enterprises were in classes with zero or unknown number of employees, while 17% of enterprises were in classes with 10 and more employees.

This distribution of enterprises by classes of employees depended upon the criteria used for extracting the frames from SBR, regarding number of employees.

Table 3a. Enterprises in the SORS sampling frames during 2013

Number of frames	Enterprises included in sampling frames by number of employees						
	Total	Unknown	0	1-9	10-49	50-249	250+
Total	**105332**	**1844**	**8396**	**77196**	**13048**	**4040**	**808**
1	42617	1202	6287	32824	2213	80	11
2	20036	395	1683	15381	637	1683	257
3	12038	206	330	10759	446	222	75
4	13588	35	91	12346	819	240	57
5	6583	6	5	4937	731	706	198
6	2520	0	0	942	724	707	147
7	2353	0	0	7	2079	243	24
8	4580	0	0	0	4400	148	32
9	605	0	0	0	587	11	7
10	410	0	0	0	410	0	0
11	2	0	0	0	2	0	0

Table 3b shows that 40.5% of enterprises were in the frame for one survey. Broken down by category, even 75% of enterprises with 0 employees were in the frame for one survey and 65% of enterprises with unknown number of employees. For enterprises with more employees, this percentage decreased from 42.5% for enterprise with 1 – 9 employees to 1.4% for enterprises with 250 and more employees.

On the other hand, 40.5% enterprises were in survey frames for three and more surveys. Within the classes, the highest shares of enterprises in three and more survey frames were in the classes with 10 – 49, 250 and more, and 50 – 249 employees: 78%, 67% and 56%, respectively. Even 34% of enterprises in the class 10 – 49 employees were in eight survey frames, which is more than the corresponding percentages of larger enterprises. This is due the fact that some of the surveys based on random samples have as target population only small enterprises.

Table 3b. Enterprises in the SORS sampling frames during 2013 (%)

Number of frames	Enterprises included in sampling frames by number of employees (%)						
	Total	Unknown	0	1-9	10-49	50-249	250+
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
1	40.5	65.2	74.9	42.5	17.0	2.0	1.4
2	19.0	21.4	20.0	19.9	4.9	41.7	31.8
3	11.4	11.2	3.9	13.9	3.4	5.5	9.3
4	12.9	1.9	1.1	16.0	6.3	5.9	7.1
5	6.2	0.3	0.1	6.4	5.6	17.5	24.5
6	2.4	0.0	0.0	1.2	5.5	17.5	18.2
7	2.2	0.0	0.0	0.0	15.9	6.0	3.0
8	4.3	0.0	0.0	0.0	33.7	3.7	4.0
9	0.6	0.0	0.0	0.0	4.5	0.3	0.9
10	0.4	0.0	0.0	0.0	3.1	0.0	0.0
11	0.0	0.0	0.0	0.0	0.0	0.0	0.0

If an enterprise was in several survey frames, it did not necessarily mean that it was overburdened, but it had a greater chance of increased burden if it was an enterprise with a large number of employees.

Tables 4a and 4b show the burden of enterprises from random sample surveys at SORS in 2013. Most of the sampled enterprises (71%) were included in only one survey, so they perceived limited statistical burden. The burden can be considered significant for about 13% of enterprises that were involved in at least three samples.

The burden of small enterprises with unknown number of employees, 0 employees and 1 – 9 employees across surveys was successively managed, as they were mostly involved in one survey (more than 91% of them). A small percentage of these enterprises (up to 0.7%) were in three or more surveys, probably because they belonged to a small domain of interest.

For enterprises with more employees, the burden from surveys was unavoidable, so that 65% of enterprises with 250 and more employees were included in samples of three or more surveys, while 23% were surveyed even in five surveys and 12% in six surveys.

Table 4a. Enterprises burdened from SORS random sample surveys in 2013

Number of frames	Enterprises included in random samples by number of employees						
	Total	Unknown	0	1-9	10-49	50-249	250+
Total	**21664**	**376**	**1304**	**10051**	**6693**	**2435**	**805**
1	15404	361	1280	9223	3619	913	8
2	3416	14	23	761	1784	561	273
3	1752	1	1	67	1008	564	111
4	570	0	0	0	219	260	91
5	333	0	0	0	52	96	185
6	144	0	0	0	10	36	98
7	25	0	0	0	0	4	21
8	16	0	0	0	1	1	14
9	4	0	0	0	0	0	4

Table 4b. Enterprises burdened from SORS random sample surveys in 2013 (%)

Number of frames	Enterprises included in random samples by number of employees (%)						
	Total	Unknown	0	1-9	10-49	50-249	250+
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
1	71.1	96.0	98.2	91.8	54.1	37.5	1.0
2	15.8	3.7	1.8	7.6	26.7	23.0	33.9
3	8.1	0.3	0.1	0.7	15.1	23.2	13.8
4	2.6	0.0	0.0	0.0	3.3	10.7	11.3
5	1.5	0.0	0.0	0.0	0.8	3.9	23.0
6	0.7	0.0	0.0	0.0	0.1	1.5	12.2
7	0.1	0.0	0.0	0.0	0.0	0.2	2.6
8	0.1	0.0	0.0	0.0	0.0	0.0	1.7
9	0.0	0.0	0.0	0.0	0.0	0.0	0.5

In sub-annual surveys, enterprises need to provide data to statistics more than once in a year. These surveys may burden enterprises, even if they are in the sample for only one survey.

At SORS, out of 21 surveys conducted on random samples in 2013 (Table 2), 15 surveys were sub-annual. There were six monthly surveys but they, in spite of the frequency of demand for responsiveness, did not present high burden with respect to the amount and complexity of required data.

By the implemented method of coordination, with rotation of one rotation group annually and with census units in most surveys, samples of the same survey are expected to have large overlap in successive years. The amount of overlap of samples depends also upon births and deaths in the business population and changes in the survey designs.

For example, Table 5 contains data on number of units in the population, sampling frame and in the sample of the Quarterly structural business survey (SBS03) in 2013 and in 2014, while Table 6 presents number of units of the survey population, sampling frame and the sample that are common, as well as those that are only in the survey 2013 or 2014.

Table 5. Number of units in the survey population, sampling frame and sample according to number of employees, SBS03 in 2013 and 2014

	Number of units in 2013				Number of units in 2014			
	Population	Frame	Sample	Census	Population	Frame	Sample	Census
Total	**80082**	**29318**	**2865**	**1524**	**82688**	**27356**	**2802**	**1459**
< 50 employees	77365	26657	1805	851	80039	24766	1696	705
50 and more employees	2717	2661	1060	673	2649	2590	1106	754

Table 6. Number of units in the SBS03 survey population, sampling frame and sample, SBS03 in 2013 and 2014: in common units and those that are in only one of those surveys

	Number of units				Number of units (%)			
	Total	In 2013 and 2014	Only in 2013	Only in 2014	Total	In 2013 and 2014	Only in 2013	Only in 2014
Population	89229	73541	6541	9147	100	82.4	7.3	10.3
Frame	33071	23604	5714	3753	100	71.4	17.3	11.3
Sample	3714	1953	912	849	100	52.6	24.6	22.9
Census	1923	1060	464	399	100	55.1	24.1	20.7

In 2014, the method of determination of census units has been slightly changed, compared to the one in 2013. As a result, the sample size has been reduced by 2% while increasing the planned precision of estimates for auxiliary variables totals (turnover and number of employees) for the main planned domains of interest[1]. The samples 2013 and 2014 have in common 53% of units and 55% of census common units.

8. Concluding remarks and suggestions for further work

Implementation of sample coordination has started at SORS in 2013. For now, the coordination system covers all business surveys which are based on random samples.

Starting from 2005, all enterprises from SBR that were included in at least one sampling frame have been assigned permanent random numbers from the uniform distribution on the interval $[0,1]$.

All random surveys of enterprises have been grouped in four blocks according to their starting point for sample selection since 2013. Selection of samples is made to the right of the starting point.

[1] Sample allocation by strata was obtained using multivariate and multi-domain procedure (Bethel, 1989, ISTAT)

Also, according to their permanent random numbers (PRN), enterprises have been randomly grouped into five rotation groups. Once a year, PRN numbers of one rotation group are shifted by 0.1 for all enterprises that existed the last time the rotation was performed.

Simple random samples or Pareto πps samples are selected using JALES technique.

The implemented method enables control of sample overlap across surveys and in repeated conduct of the same survey. However, there is still a lot of work to be done.

For further development of coordination of samples and frame populations, cooperation among the departments of economic statistics, Business register and the Department for sampling methodology is necessary.

Among the issues that should be considered are the following:

1. Is the implemented method of coordination adequate with respect to:

- number of survey blocks, according to the starting point for sample selection;
- number of rotation groups for enterprises and
- annual shift of 0.1 for PRN numbers of one rotation group.

2. The system of coordination has been implemented to random surveys of enterprises from SBR. If for some surveys samples should be selected from other units of SBR (Legal Units, Kind of Activity Units, Local Units or Local Kind of Activity Units), then they should be included in the system of survey and frame coordination.

3. The use of the same version of SBR for extraction of sampling frames for different surveys is necessary for production of comparable statistics. It is advisable to use versions of SBR after its annual updating with data on turnover or number of employees from administrative sources. For annual surveys, it is most appropriate to use the version of SPR from the end of the reference year (for example November), while for the short-term statistics the most acceptable is the version of SBR from the beginning of the year (for example March) and from August of the current year.

The same version, frozen versions of SBR, should also be used for construction of sampling frames for all surveys, not only for those that are based on random samples. The use the same version of SBR for different surveys is termed as coordination of survey frames.

Frozen version of SBR should contain identification data and auxiliary variable values that are needed for sample designs.

The address details, telephone, email, contact, etc. are not needed in the frozen version. They should be taken from the current state of SBR.

4. Crucial data and meta data should be stored for all surveys in data bases. Data should be stored with respect to the target group of interest.

For departments dealing with business surveys:

- Frozen versions of SBR including: ID of the unit, NACE Rev. 2 code, territory code[2], size measures – number of employees and turnover and date of creation of the SBR frozen version;

For Department for sampling methodology:

- Survey frames including: ID of the survey, ID of the unit, stratum information, indicator whether the unit was in the sample, date of creation of the frame;

For departments dealing with business surveys and the Department for sampling methodology:

- Survey samples containing: ID of the survey sample, ID of the unit, response code, survey reference date;

[2] Code of the settlement

For SBR department, the departments dealing with business surveys and Department for sampling methodology:

- Survey sample units with changes including: date of change, ID of the survey that is the source of information, ID of the unit, field information on the economic activity (if changed), field information on territory code (if changed).

  This data set should also include auxiliary variables with information whether these changes were conducted in SBR and date of updating;

- Survey sample units that experienced transformation (splitting, merging, change of ID code, newly opened but large): date of transformation; ID of the survey sample that is a source of information; code of transformation; for old and new unit data on: ID, stratum information, NACE Rev. 2 code, territory code, size measure; notion whether SBR was updated and if so, date of updating.

5. Data from surveys (closing of an enterprise, changes in activity or territory) should be used to evaluate the quality of SBR.

Data from surveys concerning census strata (completely enumerated) can be used to update the SBR information when samples are coordinated in time.

Census units' changes can be also used for updating the frames and the sample designs of surveys, but these changes should be conducted in all surveys that cover them as the sample census units.

Data from surveys of other units, if used for updating SBR, could introduce bias in the subsequent estimates. One way of using available survey information for updating SBR and have small bias in results is to use data for updating from all surveys at the same time. For example, updating could be conducted at the time of annual updating of SBR with administrative data.

6. The system of sample and frame coordination plays important role in measuring and reducing the enterprise burden from surveys. Survey coordination alone does not solve the burden of enterprises from surveys, particularly for medium and large enterprises and for those that belong to small domains of interest.

There are other ways of reducing the burden of enterprises from surveys. For example the burden can be reduced by:

- improving survey methodologies and removing information redundancies, when possible (collect only necessary information).
- providing alternative modes of collection of data (paper, electronic via internet, telephone and etc.) so that the enterprise can chose the one that suits it best.
- integrating surveys of the same periodicity and use of personalized questionnaires – this is especially important for large enterprises. If they are in the population of interest of a survey, then they are the same time as census units in the sample.
- promoting greater use of administrative sources of data for statistical purposes.
- strengthening the cooperation with enterprises – providing them the information concerning the use of the collected data, as well as how will they benefit if they participate in surveys.

References

Bethel, J. Sample allocation in multivariate surveys. *Survey methodology*, 15, pp 47-57 (1989).

Biggeri Luigi. *Tools for reduction of response burden in Italy: Respondents database and software for coordination of samples.*

Lindblom, A., Teterukovsky, A. *Coordination of Stratified Pareto πps Samples and Simple Random Samples at Statistics Sweden*, Papers presented at the ICES-III, June 18 – 21, Montreal, Quebec, Canada (2007).

Ohlsson, E. SAMU – The System for Co-ordination of Samples from the Business Register at Statistics Sweden. *R&D Report*, Statistics Sweden, 18 (1992).

Rosén, B. On Sampling with Probability Proportional to Size, *Journal of Statistical Planning and Inference*, 62, pp 159-191 (1997).

SAMU – The system for co-ordination of frame populations and samples from the Business Register in Sweden Department of Economic Statistics, Statistics Sweden (2003).

Appendix 1

Pareto πps is a specific probability proportional to size sample without replacement (A. Lindblom and A. Teterukovsky, 2007) in case when each unit from the sampling frame has a size measure $x_k > 0$. This sampling design makes sense when the size measure is positively correlated with the values of the study variable(s). Probability of inclusion of a unit k in a sample of size n from the population of size N by this sample design is equal:

$$\lambda_k = \frac{nx_k}{\sum_{j=1}^{N} x_j},\ k = 1,..., N \tag{1}$$

where $\lambda_k < 1$. If for some units the value of $\lambda_k \geq 1$, they are put into a census stratum and Pareto is applied to the remaining units that satisfy the condition that their probability of inclusion is less than 1. After removing m census units from the frame, inclusion probabilities λ_k are recalculated for the remaining units:

$$\lambda_k = \frac{(n-m)\cdot x_k}{\sum_{j=1}^{N-m} x_j},\ k = 1,..., N-m \tag{2}$$

Selection of the Pareto πps sample with the help of the associated permanent random numbers to the units of the frame is in the following way:

- For each non census unit from the frame calculate the value of the ranking variable q_k:

$$q_k = \frac{u_k(1-\lambda_k)}{\lambda_k(1-u_k)},\ k = 1,..., N-m \tag{3}$$

  where u_k is the assigned permanent number to unit k.

- The smallest $n-m$ units, according to q_k, form the desired sample (start point 0).

This method of selection ensures that actual inclusion probabilities π_k are approximately equal to λ_k, $\pi_k \approx \lambda_k$ (Rosen, 1997).

Similar to the case of simple random sampling, this method of selection is easily generalized to stratified Pareto πps.

Stratified Pareto πps sample can be selected using JALES technique and with surveys put in blocks. In order to be able to select the Pareto πps, irrespective of the start point a, assigned permanent number of each unit k is temporarily transformed, for sampling direction right, in the following way:

$$\begin{aligned} z_k &= u_k - a, \text{ if } u_k \geq a \\ z_k &= 1 + u_k - a, \text{ if } u_k < a \end{aligned} \tag{4}$$

The transformed permanent random numbers can be used for selection of a Pareto πps and sequential simple random sample from start point 0.

Appendix 2

Enterprises in surveys based on random samples according to number of frame populations and samples in which they are included, 2013

Number of times included			In sample											
			Total	1	2	3	4	5	6	7	8	9	10	11
In frame population	Total	**Total**	**21664**	**15404**	**3416**	**1752**	**570**	**333**	**144**	**25**	**16**	**4**	**0**	**0**
		1	**1981**	1981	0	0	0	0	0	0	0	0	0	0
		2	**3145**	2800	345	0	0	0	0	0	0	0	0	0
		3	**3152**	2705	327	120	0	0	0	0	0	0	0	0
		4	**3526**	2820	542	117	47	0	0	0	0	0	0	0
		5	**3011**	1956	494	273	150	138	0	0	0	0	0	0
		6	**1723**	779	381	285	108	82	88	0	0	0	0	0
		7	**1619**	686	422	325	94	45	33	14	0	0	0	0
		8	**2626**	1293	638	484	110	58	19	9	15	0	0	0
		9	**515**	223	145	92	39	7	3	2	0	4	0	0
		10	**364**	160	122	56	22	3	1	0	0	0	0	0
		11	**2**	1	0	0	0	0	0	0	1	0	0	0
	Unknown number of employees	**Total**	**376**	**361**	**14**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
		1	**213**	213	0	0	0	0	0	0	0	0	0	0
		2	**96**	91	5	0	0	0	0	0	0	0	0	0
		3	**46**	42	3	1	0	0	0	0	0	0	0	0
		4	**16**	11	5	0	0	0	0	0	0	0	0	0
		5	**5**	4	1	0	0	0	0	0	0	0	0	0
		6	**0**	0	0	0	0	0	0	0	0	0	0	0
		7	**0**	0	0	0	0	0	0	0	0	0	0	0
		8	**0**	0	0	0	0	0	0	0	0	0	0	0
		9	**0**	0	0	0	0	0	0	0	0	0	0	0
		10	**0**	0	0	0	0	0	0	0	0	0	0	0
		11	**0**	0	0	0	0	0	0	0	0	0	0	0
	0 employees	**Total**	**1304**	**1280**	**23**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
		1	**747**	747	0	0	0	0	0	0	0	0	0	0
		2	**455**	444	11	0	0	0	0	0	0	0	0	0
		3	**85**	73	11	1	0	0	0	0	0	0	0	0
		4	**17**	16	1	0	0	0	0	0	0	0	0	0
		5	**0**	0	0	0	0	0	0	0	0	0	0	0
		6	**0**	0	0	0	0	0	0	0	0	0	0	0
		7	**0**	0	0	0	0	0	0	0	0	0	0	0
		8	**0**	0	0	0	0	0	0	0	0	0	0	0
		9	**0**	0	0	0	0	0	0	0	0	0	0	0
		10	**0**	0	0	0	0	0	0	0	0	0	0	0
		11	**0**	0	0	0	0	0	0	0	0	0	0	0

Enterprises in surveys based on random samples according to number of frame populations and samples in which they are included, 2013 (continued)

Number of times included			In sample											
			Total	1	2	3	4	5	6	7	8	9	10	11
In frame population	1 – 9 employees	**Total**	**10051**	**9223**	**761**	**67**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
		1	**717**	717	0	0	0	0	0	0	0	0	0	0
		2	**1660**	1636	24	0	0	0	0	0	0	0	0	0
		3	**2573**	2398	171	4	0	0	0	0	0	0	0	0
		4	**2801**	2498	294	9	0	0	0	0	0	0	0	0
		5	**1800**	1590	177	33	0	0	0	0	0	0	0	0
		6	**498**	383	95	20	0	0	0	0	0	0	0	0
		7	**2**	1	0	1	0	0	0	0	0	0	0	0
		8	**0**	0	0	0	0	0	0	0	0	0	0	0
		9	**0**	0	0	0	0	0	0	0	0	0	0	0
		10	**0**	0	0	0	0	0	0	0	0	0	0	0
		11	**0**	0	0	0	0	0	0	0	0	0	0	0
	10 – 49 employees	**Total**	**6693**	**3619**	**1784**	**1008**	**219**	**52**	**10**	**0**	**1**	**0**	**0**	**0**
		1	**257**	257	0	0	0	0	0	0	0	0	0	0
		2	**180**	161	19	0	0	0	0	0	0	0	0	0
		3	**226**	139	73	14	0	0	0	0	0	0	0	0
		4	**436**	250	158	26	2	0	0	0	0	0	0	0
		5	**459**	238	158	48	15	0	0	0	0	0	0	0
		6	**458**	246	115	76	18	3	0	0	0	0	0	0
		7	**1363**	657	392	251	45	16	2	0	0	0	0	0
		8	**2450**	1289	605	447	79	25	5	0	0	0	0	0
		9	**498**	221	142	90	38	5	2	0	0	0	0	0
		10	**364**	160	122	56	22	3	1	0	0	0	0	0
		11	**2**	1	0	0	0	0	0	0	1	0	0	0
	50 – 249 employees	**Total**	**2435**	**913**	**561**	**564**	**260**	**96**	**36**	**4**	**1**	**0**	**0**	**0**
		1	**39**	39	0	0	0	0	0	0	0	0	0	0
		2	**497**	468	29	0	0	0	0	0	0	0	0	0
		3	**147**	53	62	32	0	0	0	0	0	0	0	0
		4	**199**	45	82	64	8	0	0	0	0	0	0	0
		5	**549**	124	151	168	90	16	0	0	0	0	0	0
		6	**620**	150	171	189	81	22	7	0	0	0	0	0
		7	**230**	28	30	73	49	24	22	4	0	0	0	0
		8	**144**	4	33	36	31	32	7	0	1	0	0	0
		9	**10**	2	3	2	1	2	0	0	0	0	0	0
		10	**0**	0	0	0	0	0	0	0	0	0	0	0
		11	**0**	0	0	0	0	0	0	0	0	0	0	0

Enterprises in surveys based on random samples according to number of frame populations and samples in which they are included, 2013 (continued)

Number of times included			In sample										
			Total	1	2	3	4	5	6	7	8	9	10	11
In frame population	250 and more employees	**Total**	**805**	**8**	**273**	**111**	**91**	**185**	**98**	**21**	**14**	**4**	**0**	**0**
		1	**8**	8	0	0	0	0	0	0	0	0	0	0
		2	**257**	0	257	0	0	0	0	0	0	0	0	0
		3	**75**	0	7	68	0	0	0	0	0	0	0	0
		4	**57**	0	2	18	37	0	0	0	0	0	0	0
		5	**198**	0	7	24	45	122	0	0	0	0	0	0
		6	**147**	0	0	0	9	57	81	0	0	0	0	0
		7	**24**	0	0	0	0	5	9	10	0	0	0	0
		8	**32**	0	0	1	0	1	7	9	14	0	0	0
		9	**7**	0	0	0	0	0	1	2	0	4	0	0
		10	**0**	0	0	0	0	0	0	0	0	0	0	0
		11	**0**	0	0	0	0	0	0	0	0	0	0	0

TEXT: **Olga Melovski Trpinac, Marija Ninić and Marija Panović**

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд
311 (497.11)

РАДНИ документ / Република Србија - Републички завод за статистику = Working Paper/ Republic of Serbia – Statistical Office; одговара Драган Вукмировић.
- 1965, бр. 1- . - Београд (Милана Ракића 5) : Републички завод за статистику Србије, 1965- (Београд : Републички завод за статистику Србије). - 27 cm

ISSN 1820-0141 = Радни документ - Република Србија. Републички завод за статистику

COBISS.SR-ID 59835916

www.stat.gov.rs

Published and printed by
Statistical Office of the Republic of Serbia, Belgrade, 5, Milana Rakica St
Phone: +381 2412-922; Fax: +381 2411-260
Number of pages: 23 ● Circulation: 20