

**Economic and Social Council**Distr.: General  
12 July 2017Original: English  
English and Russian only**Economic Commission for Europe**

## Conference of European Statisticians

**Group of Experts on Population and Housing Censuses****Nineteenth Meeting**

Geneva, 4–6 October 2017

Item 2 of the provisional agenda

**Innovations in census methodology and technology, and results of testing****Administrative data and sample surveys' data usage for  
determination of economic activity of population in register  
based Population and Housing Census in Latvia****Note by the Central Statistical Bureau of Latvia<sup>1</sup>***Summary*

Population and Housing Census is a unique source of socio-demographic and economic characteristics of population. The next Census will be held in 2021 in Latvia and carried out based on information obtained from administrative data sources, as well as regular statistical sample surveys, if necessary.

The Central Statistical Bureau of Latvia (CSB) has started the scheduled research and analysis of the availability of administrative data by assessing the quality of incoming administrative data, as well as developing administrative data processing methodologies and criteria for assessing the quality of data in the field of economic activity of population of Latvia.

The document presents the main results of the feasibility study of the usage of administrative data and information from statistical sample surveys to obtain data for the Census section on the economic activity of population. Furthermore, information on economic activity of Latvian population from Labour Force Survey (LFS) are compared with corresponding data from administrative data sources. Comparison of Latvian employment definitions is made with the International Labour Organisation (ILO)

<sup>1</sup> Prepared by Peteris Vegis and Anna Klusa.



definition and main conclusions, as well as solutions for future work.

## I. Introduction

1. The 2021 Census in Latvia will be carried out based on information obtained from administrative data sources, as well as regular statistical sample surveys, if necessary. Work starts with the aim to study the availability of administrative data by assessing the quality of incoming data, as well as to develop administrative data processing methodologies and criteria for assessing the quality of data processing. A feasibility study was organized in the CSB of Latvia on the use of administrative data and information from statistical sample surveys to obtain data for the Census Programme part on the economic characteristics of population.

2. Special attention was also provided to the feasibility studies of the "Social Statistics Data Warehouse" (SSDW) containing data from administrative sources to be used for the 2021 Census. The main requirements which could be introduced in the operation of the SSDW in the future with the aim to ensure optimal combination of information from various administrative data sources were identified.

3. The following Census core topics according to the *UNECE Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing*<sup>2</sup> were studied and analysed current activity status, main job, occupation, industry (branch of economic activity) and status in employment.

4. Initial assessment of the availability and quality of administrative data in the databases of the State Revenue Service (hereinafter referred to as SRS), State Social Insurance Agency (SSIA), State Employment Agency (SEA), National Education Information System (NEIS), as well as data received from Higher Education Institutions (HEI) in comparison with the data from the LFS and results of the 2011 Census was performed. In addition, some research on Farm Structure Survey (hereinafter referred to as FSS) data and data from the Rural Support Service (hereinafter referred to as RSS) and the Agricultural Data Centre (hereinafter referred to as ADC) was done.

## II. Determination of the current activity status

5. As defined by the European Commission Implementing Regulation (EU) No 2017/543<sup>3</sup> of 22 March 2017, laying down rules for the application of Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns, "Current activity status" is the current relationship of a person to economic activity, based on a reference period of one week, which may be either a specified, recent, fixed, calendar week, or the last complete calendar week, or the last seven days prior to enumeration.

6. The "currently economically active population" comprises all persons who fulfil the requirements for inclusion among the employed or the unemployed.

7. To carry out the scheduled research and analysis of the employed and unemployed persons (described below in the respective sections), it was decided to combine the data

---

<sup>2</sup> [https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41\\_EN.pdf](https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41_EN.pdf)

<sup>3</sup> <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0543&from=NE>

from administrative data sources in accordance with the situation present on 1 January 2015. A comparison was made with the relevant data from the LFS (see Annex 1) during the same time period.

8. Only those employed persons who lives in private households were selected from administrative data sources for analysis to be comparable with the data of the LFS estimate.

9. 9. Data necessary for the research of the economically active and not active population were available from several administrative data sources - information on taxpayers collected by the SRS, information on farm owners/users from the CSB Statistical Farm Register (hereinafter referred to as SFR), data from the SEA unemployed database, etc.

10. The research carried out formed a basis for the conclusion that the information included in administrative data sources can be used to acquire information on the economical activity status of persons necessary for the Programme of the 2021 Census.

11. Regarding the economically non active population, it was concluded that the pension recipients from administrative data sources are defined within the limits of the confidence interval and no significant changes are required in the method applied.

12. Work on the improvement of methodology was continued after the completion of the Eurostat Grant project to determine the number of employers, self-employed persons, income from capital recipients on base of administrative data on 1 January 2016.

### III. Determination of the economically active population

13. The CSB has access to administrative data sources – collected information of the SRS on tax payers –based on an interdepartmental agreement for acquiring information on the economical characteristics of the population foreseen in the Programme of the Census. Part of the required information (about farm owners/users) may be also found in the SFR of the CSB.

14. The indicated data sources allows acquiring information on economic activity status necessary for the population and housing census, which generally corresponds with the ILO definitions in this case. However, it should be noted that some differences were identified between the criteria used by the SRS and the employment definition of the ILO:

(a) The SRS data base also contains information on other types of income of a person, not only one's salary, as the main duty of the SRS is tax administration. Other types of income include, for example, winnings from a lottery, income from capital, etc.;

(b) Information on working hours is not available for all employed persons, since they are not recorded for certain employees;

(c) The database also includes persons under working age, as the Labour Law of the Republic of Latvia<sup>4</sup> allows to employ children from the age of 13 (with parents permitting the children to work outside of school hours, doing light work not harmful to the safety, health, morals and development of the child. The jobs in which children may be employed are specified in the regulations of the Cabinet of Ministers of the Republic of Latvia);

(d) It is not possible to determine whether a person has actually worked during the reference week for a part of employed persons (reference period in accordance with the

<sup>4</sup> [www.vvc.gov.lv/export/sites/default/docs/LRTA/Citi/Labour\\_Law.doc](http://www.vvc.gov.lv/export/sites/default/docs/LRTA/Citi/Labour_Law.doc)

recommendations of the ILO) as data for them are available only on quarterly or annual earnings;

(e) The database includes information on persons who have already ceased their working relationships during the reference period, as the SRS data reflect all tax payments, but they may also be made after the termination of employment relations. When the status of employed person is determined, the termination date of employment relations should be considered;

(f) Data on sickness benefits are collected with a long-time lag.

15. Data compiled by the RSS contain information on land owners, whereas the data of the Integrated Administration and Control System (hereinafter referred to as IACS) include details on holders and owners of herds, applying to employees only partially.

#### **A. Methodology for determining employment based on administrative data**

16. Only data related to remuneration for work are selected from tables that also contain information on other types of income other than wages and salaries.

17. Information from separate data tables of SRS is combined by using personal identification number (hereinafter referred to as PIN) as the identifier. A person is employed if at least one of the tables reflects data on the respective person's hours worked, income, as well as a periodical leave, with maintaining one's job.

18. Residents with their place of employment located abroad are listed in the following data tables:

(a) Income obtained abroad by a natural person (resident);

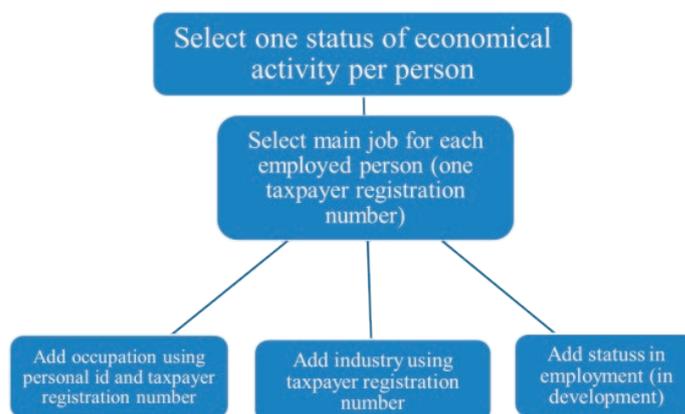
(b) Income obtained abroad by a natural person (seafarer) who is employed on a ship used for international transport.

#### **B. Methodology for determining the main job**

19. According to Paragraphs 521-526 of the *“Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing”*, individuals should be classified in correspondence with such descriptive variables as “occupation”, “industry”, “status in employment” only in compliance with their relationship with the job (main job) (see Figure 1).

Figure 1.

**Action plan of creating employment statistics for the population and housing census needs**



20. The analysis carried out during the initial project shows that approximately 80% of Latvian employees earn their income only from one job, while 20% are employed in several jobs. For it to be possible to classify which of these several jobs is one's main job, based on which a person's economic activity, occupation and status in employment is established according to the ILO definition, the hours worked or the average monthly salary were compared between them (the main job must be determined in accordance with the time spent on the job or, if that is not possible, according to the income received).

21. According to the information available in administrative data, the main job is:

(a) The one where the most hours have been worked. In cases when the number of hours worked were known for some jobs but not for others, the main job was the one in which the number of working hours exceeded half of the normal monthly working hours (it was 75 hours in December 2015). It was the main job even if the income received from it was smaller than in another job, regarding which no information on working hours was available. If all jobs had an identical number of working hours or if there was a situation when the number of working hours was not known for certain jobs and it was less than 75 hours for others, the main job was the one with the largest income;

(b) In cases when both the number of hours worked and the income were the same, the main job was determined randomly. However, it should be noted that there were only a few cases;

22. The FSS does not indicate personal income, therefore jobs on the farm were selected as the main ones only in cases when there were no other jobs that could be selected as the main ones. According to the SRS and the FSS (2013), 884.4 thousand employees were identified in private households on 1 January 2015, which is approximately 3 thousand less than in the LFS estimate, so the difference is small and missing values were not found.

23. It was identified that the number of persons employed abroad indicated in administrative data sources (6 069 or 0.69% of the total number of employed persons) was significantly smaller than the one indicated in the LFS estimate (15 638 or 1.76% of the total number of employed persons). The probable reason for this could be the fact that SRS does not have enough data on employees working in countries that are outside of the European Economic Area (according to international agreements, countries of the European Economic Area provides the SRS with information on the income of residents of Latvia, but there are no such agreements with other countries), as well as other administrative obstacles

24. The economic activity status was determined on 1 January 2015, however, according to the definition, the week of Christmas/New Year should be considered for determining economically active inhabitants, which is not typical in relation to economic activity.

25. For a percentage of persons, it is not possible to determine whether they have been economically active during the reference week (reference period according to the ILO definition).
26. Considering that the LFS estimate also includes shadow economy, it is possible that too many employed persons were identified within the framework of this project and it could be necessary to adjust the methodology by selecting stricter criteria when determining whether a person is employed.
27. Starting from 2015, the SRS data on the main place of obtaining income are available – each employed person chooses it by himself or herself. It is possible to use the respective data when determining the main job, as they mostly coincide with the main job based on hours worked estimated by the Census section of CSB in accordance with the developed methodology. It should be noted, that this methodology concerns only those 20% employed persons who are employed in two or more jobs.
28. Use of data from administrative sources will speed up the processing of data and provide more precise information on the main job of employed persons on unpaid leave and maternity leave.
29. Restrictions to use the SRS data on the main place of obtaining income – those employed at micro-enterprises do not have the possibility to choose a main job.
30. The quality of the data would also be improved by reviewing the type of gain in which income is obtained, as well as specifying to what extent they relate to employment.
31. It must be evaluated in which period the economic activity of residents should be specified according to ILO definitions and considering the experience of other countries. For example, Sweden has selected November as the reference period, when there are ordinary working days without any holidays, Austria uses October as the reference month. It was decided that November will be as the reference period for the following annual rounds of data collection in Latvia.

### **C. Methodology for determining the occupation for employed persons**

32. The Occupational code of employed persons has been included in the SRS database since 1 July 2013, according to the Regulation of the Cabinet of Ministers of Latvia. Starting with the respective date, employers must submit information regarding the occupation of employees and the number of hours worked to the SRS.
33. The information mentioned above for the year 2014 is partially available to the CSB from the table “Full list of employees” of SRS. Starting with July 2015, the respective information is received from the SRS every month.
34. Overall, information on the occupation of employees corresponds with the ILO methodology and it covers about 83% of all employed persons. However, the respective data is available not for all employees:
- (a) The Regulation mentioned above prescribes that the employer does not have to provide an occupation code for persons working in micro-enterprises;
  - (b) According to the Latvian Classification of Occupations (hereinafter referred to as LPK), an occupation code does not have to be indicated for persons who are employed based only on a work-performance contract;
  - (c) According to the Labour Law of Republic of Latvia (3), the LPK is used in labour relations under an employment contract. Thus, self-employed persons also do not

have to provide information to the SRS on the occupation code according to the respective classification.

35. If the employee changes his or her occupation, the employer provides the SRS with information regarding the occupation change by indicating the new code of the occupation and the date from which the employee has started work in the new occupation.

36. Employers use LPK, which has been developed by adapting the International Standard Classification of Occupations (ISCO-08).

37. The first step in determining the occupation of an employed person was the definition of the main job (see the description above). Detailed methodology has been developed and an algorithm has been prepared for adding occupation codes to the employed persons according to the situation on 1 January 2015.

38. During the study, the information on occupations at the level of two-digit codes obtained from administrative data was compared with LFS data (employed persons living in private households in the age group 15-74; 3rd and 4th quarter of 2014 and 1st and 2nd quarter of 2015). The compliance rate of occupation codes in LFS and SRS data at the two-digit level amounts to 61.5%, which is valued as satisfactory.

39. After comparing occupation codes at two-digit level, it can be concluded that administrative data often do not contain information on those occupations that are present in the sectors with the highest rates of shadow economy in Latvia (construction, trade, taxi services, etc.). Moreover, additional research should be carried out on those employed in agriculture, considering the seasonal nature of agricultural work.

40. Since the European Commission Regulation (EU) No 2017/543 prescribes that the information on occupations of employed persons should be provided at the ISCO-08 major group level in the 2021 Census (one-digit occupation codes), the compliance of LFS data and administrative data at the respective level was also analysed.

41. The correspondence at the level of one-digit codes was higher - 68.9% (percentage of the total number of persons, the occupation of which was known in LFS and administrative data sources). According to the assessment by occupational groups, the highest level of compliance was recorded in the 5th major group of occupations (Service and sales workers) – 80.1%. This group was followed by the 8th major group (Plant and machine operators and assemblers) – 75.6% and the 2nd major group (Professionals) – 75.3%. The lowest compliance rate – 33.4% – was recorded in the 6th major group (Skilled agricultural, forestry and fishery workers).

42. Algorithm is improved using additional information from the SRS that is available to the CSB from the second half of 2015: in accordance with the additionally received data on the changes of occupations of employees (monthly data), as well as on the main job (monthly data).

43. Development of methodology for obtaining the missing information on occupation of persons employed from regular statistical surveys performed by the CSB and imputation is done in 2017. As it can be observed in the data collected within the largest CSB surveys – LFS; EU Statistics on Income and Living Conditions (hereinafter referred to as EU-SILC), the European Health Interview Survey (hereinafter referred to as EHIS) – they provided additional information on the occupations of 2.8 thousand persons on whom there was no information available in the administrative data sources.

44. Reviewing the possibility of adding extra questions to the programme of the 2020 Agricultural Census is foreseen.

45. The work on the acquisition of new administrative data sources is continued also considering the possibility to use data of various professional associations, etc.

#### **D. Methodology for determining the industry (branch of economic activity) for employed persons**

46. To determine the industry (branch of economic activity) for employed persons the ILO uses International Standard Industrial Classification of All Economic Activities of the United Nations (ISIC Rev. 4.), but in the European Union - Statistical Classification of Economic Activities in the European Community (NACE Rev. 2) is used.

47. A 4-digit code (NACE Rev. 2) of Industry (branch of economic activity) is included in the CSB Statistical Business Register (SBR). Industry (branch of economic activity) in the main job for employed persons was determined by combining the SRS data with the SBR data.

48. One of the problems is that municipalities in the SRS reports use only one general public administration activity code (NACE 8411), but various businesses are under the supervision thereof. To solve the problem, a methodology for adding appropriate NACE code for the persons employed by enterprises of local governments was developed.

49. The purpose thereof – to correct the NACE code for persons working in enterprises/institutions of local governments (schools, libraries, boiler houses, etc.) classified under the NACE code 8411 in SBR basing on the occupation thereof.

50. By evaluating the occupations and the names of workplaces, three lists of occupations were created for which:

- (a) NACE code was not added;
- (b) NACE code was changed to another;
- (c) NACE code 8411 was used.

51. According to this methodology, the NACE code 8411 may be used for 25% of persons employed in enterprises of local governments, in 33% of cases the code should be changed to another NACE code, and a special methodology for granting NACE code for 42% of the persons, whose particular occupation may be represented in various sectors should be developed. For example, the occupation "Cleaners" (code 911201) in local governments may be present in a variety of industries – as in Section N (Administrative and support service activities), O (Public administration and defence; Compulsory social security), P (Education), Q (Human health and social work activities), as well as R (Arts, entertainment and recreation).

52. During the study, the information on the industries (at section level) obtained from the administrative data was compared with the LFS data (employed persons living in private households). It was possible to use the Section code from SBR or SRS data for 99.4% of employees that had the occupation code.

53. As the Commission Implementing Regulation (EU) No 2017/543 provides that the information on industry (branch of economic activity) of employed persons in 2021 Census must be provided at NACE section level (letter headings, except for the Section Q that consists of two parts – health and social care), the compliance between the LFS and administrative data was analysed at this level. The compliance rate accounted for 72.8% that is satisfactory.

54. The highest level of compliance in terms of NACE section level (letter code), was observed in Section Q (Human health and social work activities) – 93.5%. Slightly lower level of compliance was recorded in Section K (Financial and insurance activities) – 91.9%, Section P (Education) – 90.8%. The lowest compliance rate – 45.0% – was observed in

Section N (Administrative and support service activities), followed by Section O (Public administration and defence; compulsory social security) – 45.9%.

55. Improvement of the methodology for determining industry (branch of economic activity) of persons employed in local government institutions/enterprises, the kind of economic activity of which in registers has been coded with 8411 (General public administration activities), must be continued. Methodology for imputation of missing information should be developed. Cooperation with the SRS should be continued, thus improving quality of the administrative data.

## **E. Methodology for determining the status in employment for employed persons**

56. The status in employment in the main job is determined according to the “Status in employment” topic of Commission Implementing Regulation (EU) No 2017/543. It is prescribed by this Regulation that the status in employment in the main job is broken down as follows: employees, employers, own-account workers, others, not stated, not applicable.

57. It is possible to determine the status in employment for those persons who are employed and who have a specific main job. In Latvia, the data are available from two sources:

- (a) Information on taxpayers gathered by the SRS;
- (b) CSB SFR on owners or users of farms.

58. Persons aged 15 or over who were employed but were temporarily absent due to a maternity or parental leave, unpaid leave or temporary illness or injury during the reference week) are classified under only one category, according to the status in employment that they possessed during their last employment. Persons under the age of 15, as well as those aged 15 or over who were not employed during the reference week, are classified under “Not applicable”.

59. The following slight differences were identified between the criteria used by the SRS and the ILO definitions:

(a) Commission Implementing Regulation (EU) No 2017/543 states that an “employee” is a person who works in a “paid employment” job, that is a job where the explicit or implicit contract of employment gives the incumbent a basic remuneration, which is independent of the revenue of the unit for which he/she works (this unit may be a corporation, a non-profit institution, government unit or a household). Persons in “paid employment” jobs are typically remunerated by wages and salaries, but may be paid by commission from sales, by piece rates, bonuses or in-kind payment such as food, housing or training. The SRS data do not contain information on those workers who are paid in kind;

(b) Commission Implementing Regulation (EU) No 2017/543 states that an “own-account worker” is a person who, working on his/her own account or with one or a few partners, holds a “self-employment job” and has not engaged, on a continuous basis (including the reference week), any “employees”. In turn, according to the laws and regulations of Latvia, an own-account worker’s status is granted also for those owners of micro-enterprises who employ no more than 5 employees. Therefore, we should correct status in employment for these persons.

60. Status in employment was determined according to the information from the previously listed data sources on every inhabitant of Latvia. The indicators/variables are arranged in correspondence with the respective SRS data sources in the data file. Accurate data selection for determining the status in employment in the main job was carried out

only after it was determined whether each Latvian resident at working age is employed or not, as well as what is his or her main job. At first, all persons who are employed were selected from the data array. Then data were gradually sorted by their corresponding SRS data sources.

61. Data on own-account workers were obtained from the following data sources:

(a) Income from economic activities – this could be viewed as the main source of information on own-account workers, as information on 35.5% of own-account workers was obtained from this data source;

(b) Declaration of income from economic activities for a taxpayer who pays the fixed income tax – this data source provided on 4.4% of own-account workers;

(c) Report on sums paid to individuals – this data source was used to grant the status of an own-account workers to persons who have received income at their main jobs based on the following codes:

(i) 02 – income from sales of own agricultural produce;

(ii) 03 – income from intellectual property (except income obtained as a result of inheritance);

Overall, this is the second largest data source on own-account workers, as information was acquired on 27.6% of persons with the corresponding status.

(d) Micro-enterprise tax declaration – this data source is very extensive, as it is possible to distinguish between own-account workers, employees and employers. This data source has provided information on 20.6% of own-account workers;

(e) FSS 2013 – a person is classified as an own-account worker if the information from the respective survey is acknowledged as the main source of the workplace, and if the person does not employ any other persons. This data source provided information on 8.3% of own-account workers.

62. Employees were determined by examining all the SRS data sources and by applying the principle that an employee is a person who receives income from any other person or a company. The main data sources were:

(a) Report on mandatory social insurance contributions from the work income of employees, personal income tax and state fee of the business risk for the reference month – a person is acknowledged as an employee if he or she has received income from another person or a company, and the respective person does not pay a salary to himself/herself and is not the owner of the company. Additional information from the BR was used in data processing to find out the owner of the respective company. This data source is the main source of information on employees, as information on 94.1% of them was acquired;

(b) Micro-enterprise tax declaration – this data source indicates that a person is an employee if the respective person has a main job in which there are also other employees and the respective person is not the owner of the company. To find out whether the person is or is not the owner of the main place of work (company), additional information from the BR is used (if the taxpayer registration code (hereinafter referred to as TRC) code of the main job was the code of the company, not the PIN). The respective data source is the second most extensive source of information on employees, as information was acquired on 3.7% of employees;

63. Information on employers was obtained from the following data sources:

(a) Micro-enterprise tax declaration – the data were restructured in a separate work file where all TRC codes of the main job were linked with the PINs of employees. In order to carry out the analysis, TRC codes with at least two employees were selected, and

then the owners of the companies were searched for from the respective employees. The status of an employer was given to the persons who had identical PINs and TRC codes. If the code of the main job was the same as the company's code, additional information from the BR on owners of companies was included. This data source is the main source of information on employers and it provided information on 84.1% of employers.

(b) Income from economic activities – the search for employers was carried out in a separate work file similarly as mentioned in the item above. This data source provided information on 13,9% of employers;

The data obtained were compared with the LFS data (see Annex 2).

64. When analysing the data from administrative data sources, only employees living in private households and having a specific status in employment were selected. Differences were identified among employers, as the administrative data sources included a number of employers that was three times smaller than the one indicated in the LFS. In turn, the number of employees and own-account workers indicated in administrative data sources was by 20% larger than the one indicated in the LFS.

65. All employed persons who had their main job specified have been assigned a corresponding status in employment. Thus, imputation is not necessary.

66. Further actions needed:

(a) There is a need for additional methodological work in determining the mentioned statuses and solving debatable issues in the field of status in employment;

(b) Use of additional information on the main job of employed persons from the SRS that is available to the CSB since the second half of 2015;

(c) Examining the possibility of adding extra questions to the questionnaire of the Agricultural Census 2020, for example, if a farm owner/lessee employs any other person, including data of the employed person;

(d) Assessing the procedure according to which the status in employment is granted, as well as evaluating the possibility for representatives of certain professions and industries, such as family practitioners and dentists who receive financial resources from the National Health Service for providing health care services to their patients and simultaneously employ auxiliary staff (nurses, physician assistants), to receive the status of an employer or own-account worker, as well as whether the heads of state administration institutions may be granted the status of employers or own-account workers;

(e) Improving access to the information included in the BR and comparing the information available in the BR with SRS data.

## **F. Methodology for determining the unemployed status**

67. Latvian SEA collects information on registered unemployed persons, job seekers and persons having other status. The CSB and the SEA have concluded inter-departmental agreement on the data receipt. Unemployed persons (99.8%) and job seekers (0.2%) registered with the SEA met the ILO definition of unemployed, while persons having other status did not.

68. The differences between the criteria used by ILO and definition of unemployed used in the SEA are:

(a) Only persons who have registered with the SEA are included in the data base;

(b) A maximum age limit is defined for the registered unemployed persons; in line with the Latvian legislation, the limit is set at the age at which state old-age pension is granted (62 years and 6 months in 2016, 64 years in 2021);

(c) It has not been set that person must be able to start working within a period of 2 weeks (the law stipulates that person must be ready to enter the employment relationship immediately, by a specific term has not been defined).

69. On 1 January 2015, 79.3 thousand unemployed persons and persons seeking employment were registered with the SEA –21.5 thousand or 21.3% less than in the LFS estimate and the main reasons are following:

(a) SEA data do not include unregistered unemployed persons;

(b) The maximum age of unemployed persons registered with the SEA does not exceed retirement age, but in the LFS the upper limit corresponds to 74 years. However, the number of persons of mentioned age not covered is small, since after reaching retirement age the economic activity of persons is decreasing sharply. The LFS data indicate that approximately 1–2% of the unemployed persons are in the age group over the working age.

70. Comparison of the LFS and SEA data showed that part of the unemployed persons registered with the SEA are employed (4%) or inactive (23%) according to LFS definition.

71. The priorities of Latvian economic development plan for the period till 2020 include reduction of long-term unemployment from 20 to 10 thousand persons, as well as measures to reduce the size of shadow economy. Implementation of the tasks will lead to notably smaller difference in the data collected within the framework of the LFS and the SEA data.

72. Two possible scenarios for future action should be evaluated:

(a) In 2021 solely the SEA information in the age group 15+ should be used, by including both unemployed persons and job-seekers;

(b) The necessary imputation should be made.

## **IV. Determination of the currently not economically active population**

### **A. Methodology for determining retired persons and capital income recipients**

73. The CSB has access to SSIA data that are received based on an interdepartmental agreement between the SSIA and the CSB. The objective of the operation of the SSIA is to implement the national policy in the fields of social insurance and state social benefits, as well as administer the state retirement pension and state-funded pension scheme.

74. The differences between the criteria used by the SSIA and the ILO definitions for persons outside of the labour force are negligible:

(a) Commission Implementing Regulation (EU) No 2017/543 prescribes "Pension or capital income recipients" as one of the groups of the not economically active population, without explaining the concept of "pension", while the SSIA has 44 distinct types of services and several of them includes the word "pension" in the name. According to the "Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing", where "pension" is defined as income from former employment, specific types of services were selected;

(b) It is not possible to determine the period of the receipt of services according to SSIA data (information is provided regarding the services received during the year), which results in non-compliance in the determination of the current activity status over a short reference period, as it is recommended by the ILO.

75. The current activity status of "Pension recipients" was determined in accordance with Commission Implementing Regulation (EU) No 2017/543 – in ascribing a single activity status to each currently not economically active person, priority was given to the status of "Persons below the national minimum age for economic activity" in preference to "Pension or capital income recipients" and "students" are placed after "Pension or capital income recipients".

76. A comparison with the LFS data and data from 2011 Census suggests that recipients of pensions from administrative data sources are set within the limits of the confidence interval and no significant changes in the method applied are required, including imputation.

77. Capital income recipients include not economically active persons who receive income from property, interests, investments, rents or royalties (the "Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing" Paragraph 517, Subparagraph 2). Necessary information was obtained from the SRS.

78. The data were selected from the SRS files – the annual declaration on income "Income obtained in the Republic of Latvia during the taxation year" (2014), Report on sums paid to individuals, Declaration on income from capital gains. The information under the above codes was selected on the December 2014 (if it reflects the situation the closest to the reference date).

79. The total number of persons aged 15 and over for which the status "capital income recipients" based on the SRS data was granted comprised 42 190 or 2.5%. However, exclusion of the economically active population from the number resulted in 8 862 not economically active capital income recipients or 0.5% of the persons aged 15 and over.

80. Only the answer "O. Dividends from shares, banking interest" among all the answers to the LFS questions on income/means of subsistence was found to be suitable for making comparisons within the current activity status group "Capital income recipients". It was not possible to include the answer "N. Savings, provisions" under the "Capital income recipients", since dividends from shares, banking interest should be regarded as income, whereas savings and provisions as means of subsistence. The estimate made based on the answer "O" resulted in 305 persons, moreover the confidence interval accounted for 0 that is statistically insignificant. Thus, a conclusion may be made that LFS data cannot be used in this case.

81. The study of income codes should be continued, as well as the coordination thereof with the definitions of the Census Recommendations striving to ensure more accurate selection of the data on capital income recipients.

## **B. Methodology for determining students (outside of the labour force)**

82. Ministry of Education and Science of Latvia (hereinafter referred to as MES) compiles information on the students that are registered with the MES, aged 15 and over, and who in line with the Census definition correspond to the status "student". CSB and MES have signed an inter-agency agreement on data exchange.

83. Information about university students is provided by the higher education institutions based on agreements with the largest universities on receipt of the individual

data. As of 31 December 2014, the data were available on 32 out of 59 Latvian higher education institutions. Individual information in administrative registers was available on 66 733 students, while file "Number of students in higher education institutions in the academic year 2014/2015" indicated that there were 85 881 students. Thus, it may be concluded that information is available on 78% of students.

84. Individual data provided by the MES and higher education institutions are used. As of 1 September 2014, MES had 331 794 entries (at individual level) on students aged 15 and older. Students aged 15 and older were selected. The total number of students registered at the beginning of 2015 accounted for 89 349. The estimate based on the number of persons the current activity status of which was "Not economically active" or "Unemployed" resulted in 74 330 students.

85. In turn, the records of higher education institutions contain information on 66 733 students (at personal level). The estimate based on the number of persons the current activity status of which was "Not economically active" or "Unemployed" resulted in 22 945 students. The total number of students not economically active constituted 83 013. The number accounts for 5.0% of the total population aged 15 and over, living in private households.

86. In comparison with the LFS data, the difference in the administrative data is quite large. The study of income codes should be continued, as well as the coordination thereof with the definitions of the Census Recommendations striving to ensure more accurate selection of the data on capital income recipients and, since information on 21% of students was not acquired at individual level, within the framework of the 2021 Census, the missing data will be collected either from the Register of Students that has been established this year or based on bilateral inter-agency agreements concluded with each of the higher education institutions.

87. Evaluation and analysis of the MES database quality should be continued until 2021. The first entries in the Register of Students will be made not later than in 2018. The register will serve as a base for obtaining high-quality information on highest educational level for 2021 Census information.

### **C. Methodology for determining other not economically active population**

88. In accordance with the Commission Implementing Regulation (EU) No 2017/543, homemakers will not be listed separately in the 2021 Census. The group "Others" will include all not economically active population, who have reached the working age and who are not included in any of the above-mentioned groups.

## **V. Role of the SSDW in the merging of information from administrative data sources.**

89. One of the main objectives for ensuring that the 2021 Census will be carried out successfully is the creation of a structured individual data warehouse. When starting work on the present project, improvements were also made to the SSDW system, thus ensuring the availability of data from various administrative data sources, the storage thereof in the system, data structuring, the ability to select the necessary data from different data sources at the same time by selecting the necessary data from each source, as well as the possibility to select data in accordance with specific criteria.

90. During the research/study of the economic activity indicators of the 2021 Census Programme, the SSDW was used:

- (a) To collectively review the administrative data that should be linked together (the content thereof and the number of entries);
- (b) To compare the administrative data with the LFS data;
- (c) To carry out the selection of data in accordance with different criteria;
- (d) To export data;
- (e) To save SSDW queries, thus enabling easier repeated access to the data;
- (f) To group and aggregate primary data;
- (g) To prepare reports.

91. Proposals for the further improvement of the SSDW prepared:

- (a) A need for the restructuring of the data, which would allow to see data on a specific person from a variety of data sources in one record;
- (b) The ability to create multi-column relationship between data through a number of indicators such as personal identity number and tax payer's ID number;
- (c) Creation of the module for linking the data with classifiers, which are stored in another system.

92. The main challenges/problems when carrying out the pre-treatment and analysis of data before giving the data to the users of SSDW:

- (a) Large volume of data that often exceeds the system's current technical capabilities during the processing of personal data – may be prevented by improving the SSDW;
- (b) Joining of data tables related by meaning and logics into data sources – requires in-depth content analysis of the available data in order to establish proper links between the tables of the data source;
- (c) Quality of the administrative data received, for example, inaccurate PINs – an external factor, which cannot be directly influenced by the CSB. Technical, logical and mathematical verification of compliance of person identifiers is carried out before the transfer of the data to users.

## **VI. Conclusions and recommendations for future work.**

93. SRS data are of sufficient quality to identify employed persons. More metadata information is required for it to be possible to fully assess the timeliness, coverage and comparability. Additional statistical analysis should be carried out to identify the most suitable sources if the content and format of available data is changing.

94. Employed persons in Agriculture may be identified by using the FSS. Results of FSS 2016 should be evaluated. There is a need for further research to assess the possible inclusion of additional questions in the questionnaire of the 2020 Agriculture Census.

95. There is a need for further studies on the RSS and the ADC data regarding the possibilities to use the respective data for the needs of the population and housing census.

96. Two probable future action models for the unemployed should be assessed – the sole use of SEA information for the 15+ age group in 2021, including both unemployed persons and job-seekers, or carrying out the necessary imputation to prevent essential difference compared to LFS data.

97. When concluding new inter-institutional agreements, the CSB should also include the clause on the provision of metadata, as well as timely notification about the upcoming changes within the administrative data source.
98. While developing the database necessary for the 2021 Census, information on related companies and housing should be added to the SSDW.
99. In cases when it is difficult to determine the local unit (LU), which employs a person, it could be determined according to his or her place of residence – the nearest LU to the place of residence within the municipality, or the largest one if there are several units.
100. In the case of Latvia the information on the LU is maintained only at the level of the respective unit (only the number of employees at each LU is known).
101. CSB has enough data to determine the main job by definition, but it is necessary to carry out the further improvement of methodology.
102. If the occupation and industry of the employed person cannot be determined, the mathematical statistics methods should be applied. Further in-depth studies are required.
103. It is necessary to continue studying the issues related to the employment status by providing special attention to the improvement of the methodology for determining the employers and self-employed status.
104. It is necessary to continue improving the methodology for identifying the secondary and tertiary students who are not economically active, as well as pension and capital income recipients in accordance with the implementing regulation of the European Commission (EU) No 2017/543.
105. By continuing the development and improvement of the methodology, the balance between the accuracy of statistics and the complexity of the system that produces statistics should be assessed through giving up minor improvements in accuracy if the statistical production system is made significantly more complex.
106. Improvement of the SSDW should be continued in accordance with the needs of the 2021 Census.
107. As it was started with data on 1 January 2015 and continued with data on 1 January 2016, it is necessary to develop an annual database by collecting personal data on 1 January of the respective year in accordance with the Programme of the Census and by following the changes in the laws and regulations related to the content of the information of the administrative data sources.

## Annexes

## Annex I

### Current activity status of persons aged 15 years and over living in private households

<i>Current activity status</i>	<i>Administrative data sources as of 1 January 2015</i>		<i>LFS</i>			
	<i>Number</i>	<i>%</i>	<i>Estimations</i>		<i>Accuracy of estimations</i>	
			<i>Number</i>	<i>%</i>	<i>Lower limit of 95% confidence interval</i>	<i>Upper limit of 95% confidence interval</i>
Employed person	884 423	53.0	887 509	53.3	881 526	893 493
Unemployed person	79 350	4.8	100 820	6.1	96 548	105 091
Pension recipient	403 795	24.2	406 781	24.4	403 670	409 891
Capital income recipient	8 741	0.5	x	x	x	x
Student (not economically active)	83 013	5.0	112 336	6.7	109 587	115 085
Other	209 007	12.5	157 470	9.5	152 176	162 764
<b>Total number of persons aged 15 years and over living in private households</b>	<b>1 668 329</b>	<b>100.0</b>	<b>1 664 916</b>	<b>100.0</b>	<b>1 664 916</b>	<b>1 664 916</b>
Besides:						
Persons below the national minimum age for economic activity	296 013		295 876		295 876	295 876
Total number of population living in private households	1 964 342		1 960 792		1 960 792	1 960 792

## Annex II

### Status in employment of employed persons aged 15 years and over living in private households

<i>Status in employment</i>	<i>LFS estimation for employed population aged 15–74</i>			<i>Estimation of administrative data for employed population aged 15 years and over</i>	<i>2011 Census data for employed population aged 15 years and over</i>
	<i>Estimation of four quarters</i>	<i>Lower limit of 95% confidence interval</i>	<i>Upper limit of 95% confidence interval</i>		
1 – Employee	785 500	779 525	791 476	800 576	760 905
2 – Employer	34 694	31 530	37 859	10 451	32 958
3 – Own-account worker	60 784	56 977	64 590	73 396	30 678
4 – Contributing family worker	6 531	5 020	8 041	0	1 352
Not stated	0	–	–	0	4 458
<b>Total number of employed persons</b>	<b>887 509</b>	–	–	<b>884 423</b>	<b>830 351</b>