



# Economic and Social Council

Distr.: General  
12 July 2022

Original: English

---

## Economic Commission for Europe

### Conference of European Statisticians

#### Group of Experts on Population and Housing Censuses

##### Twenty-fourth Meeting

Geneva, 21–23 September 2022

Item 5 of the provisional agenda

**Transitions in census methodology; plans, experiences and innovations**

### **Administrative register-based Census in Latvia: transition of methods, data publishing, challenges, and next steps**

**Note by Central Statistical Bureau of Latvia\***

#### *Summary*

In Latvia, the 2021 census was conducted using only personal data available from administrative data sources. Since 2012, purposeful work has been carried out to improve Latvia's largest administrative registers and to develop appropriate methods for preparing all mandatory census indicators. In 2012, a population estimation method was developed.

In cooperation with the state and local government information systems, the necessary data were identified at the level of persons and addresses (addressing codes) to ensure the preparation of indicators on households and housing.

In addition to the information available in the registers, imputation methods were used to prepare the missing values for the education and economic activity indicators.

Latvia was the first country in the European Union to publish census results. In order to promote the data and explain to data users the methodological specificities related to the transition to an administrative data census, several publicity events were organized—a data hackathon, an academic breakfast and an international conference.

Work on the census and the compilation of population indicators using administrative registers will continue and will involve developing new methods, proposing improvements to the owners of national administrative registers, and actively participating in the process of drafting the European statistics on population (ESOP) regulation.

---

\* Prepared by Sigita Purona-Sida, Baiba Zukula, Peteris Vegis, Sigita Meldere, Ruta Beinare.

*Note:* The designations employed in this document do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.



## I. Introduction

1. In 2021, for the first time in Latvia, a census based on registers was conducted. Already in 2012 a method for determining the population size was developed because the 2011 census suggested that at the beginning of the year there were 7 per cent fewer people living in Latvia than declared in the Population Register supervised by the Office of Citizenship and Migration Affairs.
2. The next step was to assess the availability, quality and reliability of the various registers and administrative data sources. Latvia has many high-quality, regularly updated registers and databases with good coverage which provide information on both persons and housing.
3. The required data for the census variables in the administrative data sources were identified at the level of persons and housing (addressing codes) to ensure the production of individual-level indicators.
4. During the preparation process, it was acknowledged that the state administrative registers cannot provide all the information necessary for the census—some groups are missing or the information is incomplete. Therefore, data were collected from non-governmental institutions and private companies.
5. No administrative register has been set up to provide data for statistical purposes, so methodological work has been done to produce all census indicators in a timely and high-quality manner.

## II. Change of method – census based on administrative registers

### A. Population and families

#### 1. Population estimation method – development and changes

6. In the 2011 census, it was established that 7 per cent (155,000) fewer people actually live in Latvia than are registered in the Population Register (since July of 2021, Register of Natural Persons) of the Office of Citizenship and Migration Affairs (OCMA).
7. Since 2012, population statistics of Latvia have been produced using a model-based approach. Every year for every person declared in the Register of Natural Persons, the probability that the person belongs to the population of usual residents of Latvia is calculated using of logistic regression model (logit model) and data from 15 administrative registers.

#### 2. Data on family and households

8. The population estimations performed by the Central Statistical Bureau (CBS) are used to determine the family nucleus. In the census, a family (family nucleus) is considered to be two or more persons living in the same dwelling who belong to one of the six family types.<sup>1</sup>
9. The main conditions for dividing the population into family nuclei are as follows:
  - (a) One person or group of persons who does not form any of the six family types is not considered a family;
  - (b) Persons living in institutional housing are not included;
  - (c) The nucleus of the family consists of a married or cohabiting couple with or without children, as well as one parent with one or more children;
  - (d) Only direct kinship between offspring and parents is taken into account. A grandparent with his or her grandchildren is not counted as a family nucleus in this context;

---

<sup>1</sup> <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1493361707846&uri=CELEX:32017R0543>.

(e) The age of the offspring is not considered. Adult men and women who are living with their parents also have the status of sons and daughters;

(f) According to the census methodology, several family nuclei can live in one household. For example, in the case of a grandmother and grandfather and one of their children with their family two family nuclei are formed.

10. The CSB has developed an algorithm for determining the types of families which uses data on Latvian permanent residents at the individual-level for persons who provide information on kinship. The family type classification scheme is available in the reference metadata for [family nucleus](#) on the Latvia Official Statistics portal.<sup>2</sup>

## B. Education and economic activities

### 1. Determination of the highest level of education

11. The highest level of education obtained is determined for each usual resident of Latvia at the beginning of the year (January 1st of the respective year). A cumulative database is used to determine the highest level of education obtained. It is based on the results of the 2011 census, adding data on the level of education acquired by the population from registers and CSB surveys in the following years. The level of education is aligned to the International Standardized Classification of Education (ISCED) 2011 before being added to the cumulative database. The quality of the administrative and other data used is being assessed.

12. When all indicators have been prepared, the quality of data, the correspondence of the acquired level of education to the person's age and to certain professions is evaluated. It is assumed that each level of education has a minimum age at which it can be obtained (table 1).

Table 1

#### Level of education and minimum age

<i>Level of education</i>	<i>Minimum age</i>
ISCED level 1. First stage of basic education	12 years
ISCED level 2. Second stage of basic education	14 years
ISCED level 3. Secondary education	17 years
ISCED level 4. Vocational education after general or vocational secondary education; vocational secondary education after general secondary education (not higher education)	18 years
ISCED level 5. First level professional higher education	19 years
ISCED level 6. Bachelor's degree or equivalent	20 years
ISCED level 7. Master's degree (including professional); second level professional higher education lasting 5 years	21 years
ISCED level 8. Doctoral degree	25 years

13. The level of education from administrative data cannot be determined, or it is not known exactly (only it is known that it is higher education) for 5.8 per cent of the Latvian population aged 15 and over (as of January 1, 2021). Given that this is a relatively small proportion, imputations have been made to replace missing values with reliable and consistent values using mathematical methods.<sup>3</sup>

<sup>2</sup> <https://stat.gov.lv/en/metadata/5744-family-nucleus>.

<sup>3</sup> <https://stat.gov.lv/en/metadata/7299-educational-attainment>.

14. Education level imputation or replacement of missing values with reliable and consistent values is necessary because on January 1st, 2021:

(a) The level of education is unknown (approximately 35 thousand persons or 2.2 per cent of the population aged 15 and over); or

(b) It is known that a person has higher education, which needs to be defined more precisely according to ISCED levels (approximately 56 thousand persons or 3.6 per cent of the population aged 15 and over).

## 2. Determination of the economic activities

15. The activity status of a person is defined for each usual resident of Latvia by merging the data from [administrative registers and CSB surveys](#). Based on the international practice, the economically active population engaged in the labour market is defined based on the reference period of November of the previous year (2020) instead of the critical moment of the Population and Housing Census. November is chosen as typical month of the year, bearing in mind that December is the closing month of the year and has a considerable number of holidays and greater fluctuations in the activity status. The data acquired are attributed to the critical moment of the census, 1 January 2021.

16. Only one activity status is defined for each person. It is done as follows:

(a) First, children that in compliance with the Section 37 of the Labour Law have not reached the national minimum age for economic activity (those aged under 14) are found; followed by classification of:

- (b) Unemployed persons;
- (c) Employed persons;
- (d) Non-working pensioners;
- (e) Capital income recipients;
- (f) Students;
- (g) Other.

17. Data on the workplace of each usual resident of Latvia in November 2020 are gathered. For persons having several jobs, one main job is defined. The main job is defined as follows:

(a) A job for which payroll tax booklet information is registered in the State Revenue Service;

(b) If there is no information about the payroll tax booklet, the main job is the one in which the higher number of hours is worked;

(c) If the number of hours worked is not known, the main job is the one in which the wage or salary is higher at the reference moment.

18. Based on the main job of the person, the following information is defined:

- (a) Occupation;
- (b) Industry;
- (c) Status in employment;
- (d) Location of the place of work.

19. The data on occupation and industry are imputed considering that administrative registers partly miss data on industry (0.7 per cent) and occupation (4.1 per cent) of the employed population.

### 3. Other benefits of calculating indicators from administrative registers for education and economic activities

20. The preparation of census indicators from administrative registers has provided an opportunity to develop other indicators of national importance. Since 2019, the Central Statistical Bureau of Latvia has been preparing data on Latvian higher education institutions and vocational education institutions including on graduate employment and income, popular study programmes, and demand for graduates in the labour market.

21. Monitoring is important for two main target groups: prospective students and universities. Research can help students make choices about particular university and study programmes, while universities can use research to improve study programmes, create new ones, as well as for marketing purposes to attract new students.

## C. Housing

22. The main sources for housing data were the State Address Register (SAR) and the Real Estate Cadastre Information System (RESCIS) of the State Land Service (SLS) and the 2011 Census.

23. Since late 2018, SLS senior experts, who are responsible for RESCIS, and CSB experts established direct communication, which improved the exchange of information and significantly increased data availability and quality.

24. To get additional data on disused housing, information about destroyed, burned or demolished residential buildings was collected from local governments. Additional information about availability of housing facilities was collected from local governments' central water supply, central sewerage, and central heating enterprises.

25. Geographic co-ordinates had been taken mainly from the RESCIS as they were available on the level of buildings and available data in general was sufficient for the 2021 Census grid.

26. During preparation for the census some algorithms and a list of conditions were developed to supplement the missing data for several housing indicators.<sup>4</sup>

27. The CSB has established a list of institutional housing (mainly for the purpose of population assessment) and this list with addresses is considered the primary source of institutional housing and, accordingly, of persons living in institutional housing.

28. Persons were recognized as homeless if they had been registered in night shelters.

29. Since 2019, housing data are added to the Social Statistics Warehouse (SSDW) annually to ensure the successful organization of a register-based Population and Housing Census in 2021.

## III. Data publishing and communication

### A. Data publishing

30. In Latvia, the Census 2021 data correspond to the situation on 1 January. The main descriptive population and migration indicators—prepared from administrative data—were the first results to be released on 28 May 2021. This included new indicators that had not previously been published: population by place of previous residence one year before Census, by year of arrival in Latvia, and by territories (neighbourhoods and densely populated areas, grid data up to 1 square kilometre and hectare). Release of the main population indicators was followed by indicators for households and families providing new detailed data on family nucleus and household size, type, and structure down to the level of municipalities.

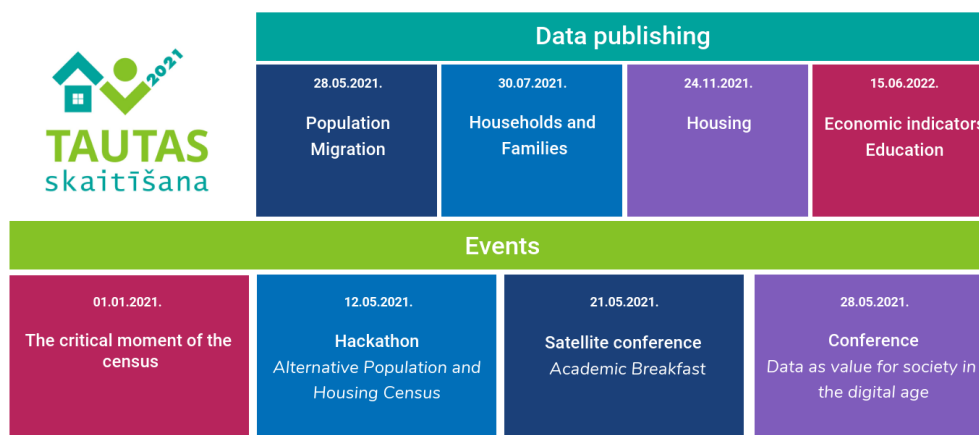
<sup>4</sup> <https://stat.gov.lv/en/metadata/10258-housing-indicators-housing-and-population-census-2021>.

31. In November 2021, housing data were published. Housing data were last provided to data users from Census 2011. The indicators included housing type, useful floor space, density, ownership, number of residents, year of construction, and facilities by territories down to one hectare in size.

32. The last data were published in June 2022 and included indicators on education and economic activity. This was the most time-consuming data collection as it required information from many taxpayers' declarations that, for the time period required for Census 2021, were available only at the end of September 2021. These indicators were available for the first time since 2011 at such detailed geographic levels. Published data included information on economic activity status, employment status, industry, occupation, place of work, and highest educational level attained. All data were published in the Official Statistics portal of Latvia.

Figure 1

### Announcement for public on main activities of Census 2021



33. By obtaining data from administrative registers, detailed demographic data, data on the level of education and employment, and separate data on housing can be published annually.

## B. Communication campaign

34. A communication campaign was organized to effectively reach different target groups.

### 1. Academics – data hackathon and academic breakfast

35. Eleven teams consisting of both Latvian university students and experienced data enthusiasts participated in hackathon. The event took place over three days and included lectures by professors from both the University of Latvia and the University of Southampton. Hackathon participants worked in teams to address specific challenges of the register-based census: How to determine whether a person's declared place of residence is also his or her actual place of residence? What factors contribute to internal migration? What social and demographic factors influence a person's desire to get an education? What is the situation with the availability of comfortable housing in Latvian cities and rural areas? How much of housing is rented in Latvia? These were just some of the issues raised by hackathon participants.

36. An academic breakfast was held to discuss with researchers a number of census topics. For the topics “demographics and the family” and “ethnic indicators and language” participants considered (a) how statistical methods can help with data and the analysis of demographic processes in the context of the new solutions for data collection in 2021 (less data collected in face-to-face interviews; register data and big data; identification and use of new data sources; detailed territorial breakdowns); and (b) which data sources should be used in the areas where data mining should be developed. For the topic of “housing indicators and



territorial aspects”, participants discussed data user needs and the key needs for territorial research. The use of census results in research and studies in public administration and local governments was also discussed at the event.

## 2. National and international data users – conference

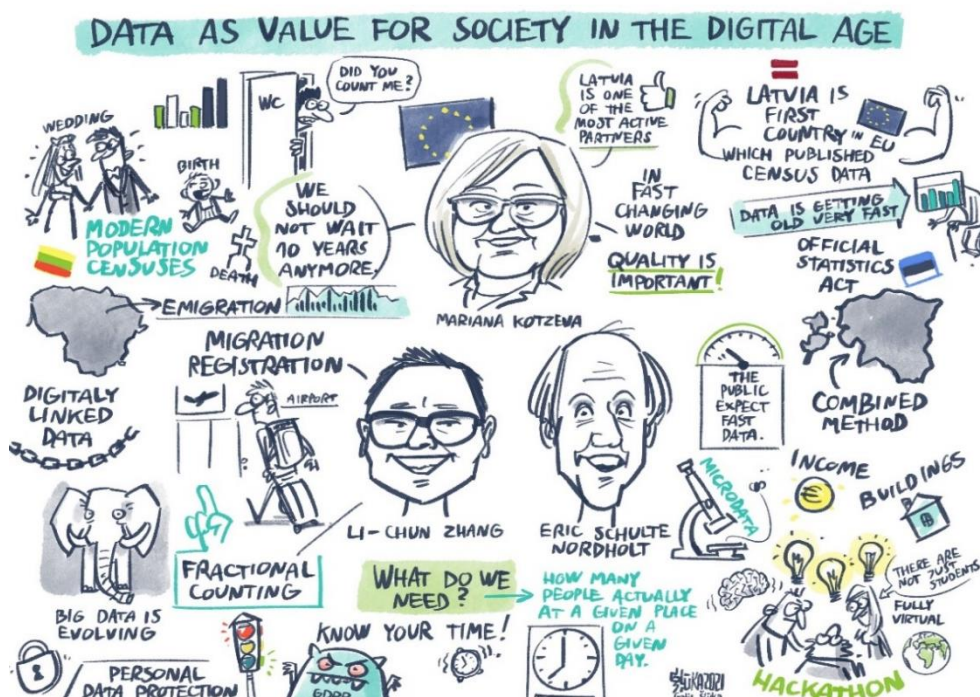
37. On the day the first census results were published, the Central Statistical Bureau of Latvia held a conference on the topic change of the paradigm and method when carrying out the Population and Housing Census 2021 in the digital age. Conference participants had the opportunity to evaluate three aspects:

- (a) Registers and the future;
- (b) Data opening and detailing;
- (c) Legal aspects, confidentiality, and protection of personal data.

38. Conference participants were invited to participate in two consecutive sessions. The first session was aimed at national data users (state and municipal institutions and media), while the second session was focused on foreign national statistical institutes and the academic community.

Figure 2

A caricature by Gatis Šļūka about the conference



## 3. Mass media and the public

39. The third group – the mass media and the public – was targeted from the beginning of 2021 with information on the transition to a census based on administrative registers and to promote the use of data. Several press releases, cartoons, infographics, and video reports were prepared.

## IV. Challenges and next steps

### A. New population method

40. Administrative data-based censuses have many advantages: data can be obtained more quickly, data acquisition is cheaper, and fewer human resources are required. However, there are also shortcomings in compiling statistics from administrative registers. No administrative

register has been set up for statistical purposes. Many registers do not use uniform IDs and classifiers, the definitions of indicators differ across registers, and data are not stored for further digital processing. In addition, administrative registers are not able to provide complete information on the entire population of the country, such as historical data on the education of the population or up-to-date information on the person's actual place of residence.

41. Work is underway to improve existing methods and develop new methods to ensure the quality of indicators obtained from administrative registers. Since 2020, the Sol-logit method has been used to provide more precise population and migration statistics. Signs of life in 12 administrative registers allowed for the assessment of the probability of being usually resident of Latvia for persons registered in the Population Register and for persons with short-term residence permits (Extended Population Register).

42. Several audit surveys (in 2015 and 2017) found that the logit model overestimates the population of Latvia. The Sol-logit model gives a result that shows that the logit model underestimates the population.

43. Research and consultation with experts have shown that there are a number of other methods for estimating the population using signs of life in administrative registers that need to be assessed and that coverage surveys should be organized to assess how precise are models.

## **B. European statistics on population (ESOP) requirements**

44. There are many challenges associated with the new ESOP regulation<sup>5</sup>, which is still under preparation. Some of the biggest challenges are the diversity of perspectives about the goal of the regulation, and about the data that are to be required from national statistical offices.

45. Latvia has shared feedback regarding changes in the regulation during the preparation phase, including:

(a) In Latvia's opinion, the ESOP process has not proceeded as usual: the Census Working Group and Census Task Force have not been included sufficiently, and objections provided during meetings of Directors of Social Statistics have, in Latvia's estimation, not been taken into account;

(b) Some population and housing data must be provided annually (starting from 2026/2027);

(c) More detailed data is needed to produce GRID maps;

(d) Eurostat plans to include in the ESOP regulation several indicators on health, energy and the environment. Given that more and more countries are moving to register-based censuses, the availability of such indicators depends largely on the availability of data in registers. Before including such indicators, it is necessary to study the availability of data and prepare methodological guidelines;

(e) The concept of "family" is out of scope for the census. Instead, the term "family nucleus" should be maintained;

(f) Latvia does not support annual housing data nor the inclusion of non-core census topics;

(g) According to the discussion at European Statistical System Committee (ESSC), equality data should not be included in ESOP;

(h) Latvia does not support the inclusion of abortions in ESOP. This indicator would better fit within the scope of health statistics.

---

<sup>5</sup> [https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12958-Data-collection-European-statistics-on-population-ESOP-\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12958-Data-collection-European-statistics-on-population-ESOP-_en).



### C. Work on improving administrative data

46. In 2021 a solely register-based census was conducted in Latvia for the first time. The census reference date was 1 January 2021. The decision to do a fully register-based census in 2021 was made in 2015 for economic (high cost of the traditional census) and social reasons (decreasing response in social surveys). In 2020, when all face-to-face (F2F) interviewers were completely stopped due to the Covid-19 pandemic, the decision to organize the Population and Housing Census using only administrative data proved to be the right one. Latvia had time to prepare, to develop a methodology for preparing indicators, and to work on improving the quality of administrative data.

47. Preparation for the 2021 census started in 2012. The CBS of Latvia studied the availability, quality and reliability of different registers and administrative data. All administrative data were evaluated annually according to the European Statistics Code of Practice and census requirements since 2017. The CSB has collaborated with register holders to ensure data exchange and coordination between institutions.

48. The CSB has concluded that the state's administrative registers alone cannot provide all the information required for the census. Therefore, data from non-governmental institutions (e.g., artist unions) and private companies (e.g., water and sewerage service companies) were gathered. More than 40 different registers served as data sources for the census.

### D. The Register of Natural Persons

49. Until 1 July 2021, the Population Register of the Office of Citizenship and Migration Affairs was the main register of the population, in which data on the basic information of each resident of Latvia were stored. The Register of Natural Persons established after that date contains information on the following groups:

- (a) Citizens and non-citizens of Latvia;
- (b) Foreigners:
  - (i) Who have received a residence permit in Latvia, a registration certificate of a European Union citizen, or a permanent residence card of a European Union citizen;
  - (ii) Who have been granted the status of stateless person in Latvia;
  - (iii) Who have submitted documents requesting the residence permit, the registration certificate, or the permanent residence card;
  - (iv) Who have legal ties with Latvia on the basis of which mutual rights and obligations are formed or have been established in the field of immovable property, commercial activity, health, taxes, benefits, and education;
  - (v) Who wish to receive a Latvian identity card (electronic identification card) in order to promote the development of economic, scientific, educational or cultural relations;
  - (vi) Who wish to receive State administration services in Latvia electronically by means of electronic identification issued by a European Union Member State, a state of the European Economic Zone, or the Swiss Confederation;
  - (vii) Who are employees of a foreign diplomatic mission accredited in Latvia, employees of a foreign consular institution, employees or officers of another international body governed by public law or its representation accredited in Latvia, as well as family members or private staff of such an employee or officer whose status in Latvia is determined by an international treaty and who has agreed to have his or her information included and updated in the Register;
  - (viii) Whose birth or marriage is performed and registered in Latvia or who repeatedly request a document confirming the registration of marriage or birth;
  - (ix) Who are asylum seekers in the Republic of Latvia.

50. Work is underway to evaluate and analyse the new data and to consider which groups should and should not be included in the usual resident population of Latvia.

## **E. Municipal information systems**

51. Latvian municipalities use a set of mutually integrated IT solutions known as the Unified Local Government System (GSP). In order to estimate the population, the CSB uses individual-level data on municipal social benefits and services from the social assistance and social services administration application (SOPA) of the GSP. Although both the CSB and policy planners could make full use of all SOPA data and they would be of high quality (complete coverage, regularly updated), it is necessary to develop a framework for uniform data processing in all municipalities in cooperation with the Ministry of Welfare.

52. An application program for the registration and circulation of municipal apartments (BRIDZIS) has been developed to track municipal housing stock and organize housing assistance and is currently used by more than half of the municipalities. In order to ensure unified computerized accounting of this information, it is necessary to regulate the inclusion and regular updating of all municipal information on municipal housing stock and housing assistance in the BRIDZIS application.

53. The regulations of the Cabinet of Ministers on the management and registration of decentralized sewerage systems include the task of registering the city's decentralized wastewater collection or treatment systems by 31 December 2021 and arranging them in accordance with regulatory requirements. The current regulation obliges local governments to establish such registers, however, the lack of uniform and systematic accounting and processing of these data does not allow for the full use of these data for monitoring and analysis. Based on the databases compiled by local governments, a state-level register should be created for local governments by initiating a unified systematic inventory of data using common classifications.

## **F. Heating equipment registers in large cities**

54. According to the Order of the Cabinet of Ministers on the Air Pollution Reduction Action Plan 2020–2030, the municipalities of Riga, Liepaja and Rezekne should collect information on the heating equipment used in households and develop a solution for collecting such information in the future. A unified system should be established in the municipalities of large cities, ensuring integration based on the classifications in use in the country.

## **G. Building Information System of the State Construction Control Bureau**

55. The work to improve the Building Information System (BIS) of the State Construction Control Bureau (SCCB) and to develop electronic versions of House and Operational files should continue. House files provide information and documentation on the ownership and technical specification of buildings. Operational files provide information on maintenance and work performed on buildings. The development of an electronic file system for the BIS will ensure a unified approach to management as well as the re-use of house file information. Work is also ongoing to improve the energy certificate module of the BIS which records information on fuel use.

## **H. State Land Service Real Estate State Cadastre Information System**

56. Appropriate amendments to the regulatory framework are required for the SLS to receive additional data on the improvement of residential buildings and groups of residential premises (apartments) from local governments and state institutions.

57. Without waiting for the cadastral survey, it must be possible for the owner to indicate information about the improvement of the dwelling. On a regular basis, for example, every

three years, the owner should receive an invitation to update the information together with the real estate tax return. It could be in the form of a survey with defined fields. In this way, it would also be possible to gather information on the measures taken to improve the facilities or the renewable energy installations installed.

58. When available, data from the SCCB BIS house files can supplement data in the RESCIS.

59. As far as possible, it is necessary to obtain additional information on the years of construction of residential buildings.

## **V. Conclusions**

60. The decision to release the census only from the administrative registers was forward-looking. Not only because it is the most cost-effective, but also because most census figures can be produced more quickly in this way. Latvia was one of the countries where the census took place according to the planned schedule despite the restrictions of the pandemic. It is planned to provide detailed and timely census indicators every year, therefore systematic work on the development of methods and improvement of the quality of registers will continue.

---