2011 POPULATION AND HOUSING CENSUS

METHODOLOGY

Ene-Margit Tiit

nr	värv1	värv 2
1	70%	20%
2	20%	80%
3	50%	50%
4	85%	15%

2011 POPULATION AND HOUSING CENSUS

METHODOLOGY

Ene-Margit Tiit

I would like to thank my good colleagues – Diana Beltadze, Kaja Sõstra, Riina Leinbock, Marika Inno, Vassili Levenko, Helen Mett, Kristi Lehto and Anu Tõnurist – for their useful recommendations and comments, which were very helpful in preparing this publication.

Ene-Margit Tiit

CONTENTS

PLANNING AND PREPARATION OF CENSUSES

Purpose of censuses

Censuses have been organised already for thousands of years. Nowadays, population and housing censuses are conducted worldwide under the aegis of the United Nations, at an average interval of ten years. Censuses provide the basis for all sorts of global forecasts and projections – the census data are used to estimate the number of people on earth and make population projections, to assess human development in all countries of the world and rank countries by the human development index, to analyse people's quality of life and resources, to compare countries and regions. To make all this possible, censuses follow a very strict set of rules (revised by the Conference of European Statisticians in 2004–2006, as is customary prior to each census round). The recommendations were developed further in cooperation between the United Nations Economic Commission for Europe (UNECE) and Eurostat (the statistical office of the European Union). These rules apply mostly to European countries and are mandatory for European Union Member States. The rules harmonise the definitions and classifications of the characteristics to be measured. The rules also establish the division of characteristics into mandatory core topics (hereinafter **core characteristics**) and optional non-core topics (**voluntary characteristics***)* that countries can choose to include in the census. Even the assessment of the quality of the census and individual characteristics is governed by international regulations (primarily adopted by Eurostat).

Censuses are organised based on the following principles (UNECE 2006):

- **individual enumeration** – the information on each enumerated person is registered individually, and the same requirement applies to dwellings;
- **simultaneity** – the collected information refers to a specific moment in time (the census moment);
- **universality** – the population and housing census has to collect information on all the persons residing/staying in the given area and on all households and dwellings in the given area; census coverage has to be checked;
- **small-area data** – the census also collects information on small geographic areas and population groups, ensuring adherence to confidentiality requirements;
- **defined periodicity**– censuses take place at regular intervals to ensure the comparability of results.

However, in any population and housing census, the national and especially the local dimension is the most important. Despite international coordination and harmonisation, each country has a unique enumeration programme that covers specific topics important in that particular country (in addition to general universal topics). Although there are very many surveys conducted in the developed world today, the census is extraordinary due to its universality. This means that a census may uncover phenomena that would remain unnoticed in other surveys. The importance of the census in enhancing national self-fulfilment cannot be overestimated – it serves as a coordinate system, underlying official statistics and all other surveys.

Thus, the purpose of the population and housing census is to determine the population composition, household and family structure, living conditions and the geographic distribution of the population. The census establishes population composition in terms of sex and age, education, ethnicity and citizenship, employment, social and occupational status, origin, language skills, and health status.

History of censuses in Estonia

The history of censuses in Estonia dates back to 1881 when the Russian Empire carried out the first census in the Estonian, Livonian and Curonian provinces. The second census in Estonia took place in 1897, as part of the first empire-wide census in Russia. The next censuses in 1922 and 1934 were organised by the Republic of Estonia. The German occupation forces conducted a census in 1941. There were four censuses during the Soviet era in 1959, 1970, 1979 and 1989. The first census after the restoration of independence took place in 2000.

The 1922, 1934 and 2000 censuses were essentially population and housing censuses; and the 1989 census also included questions about housing. In case of the censuses organised by the Republic of Estonia, each census was more detailed than the previous one. Also, where possible, the questions and answers were worded similarly to previous censuses to ensure comparability of the results.

Legal basis of the census

Today, population and housing censuses are regulated internationally by the United Nations (UN). UNECE (United Nations Economic Commission for Europe), the European Union and the European Commission have developed more

detailed regulations for Europe with requirements and recommendations for methodology, uniform mandatory output and quality analysis. There is range of guidelines where the requirements are defined and explained in detail.

Main international documents governing the organisation of population and housing censuses:

- Principles and Recommendations for Population and Housing Censuses. Revision 2. UNSD, 2008;
- Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing. UNECE, 2006;
- Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses. Official Journal of the European Union, L 218, 13 August 2008, pp. 14–20;
- Commission Regulation (EC) No 1201/2009 of 30 November 2009 implementing 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. Official Journal of the European Union, L 329, 15 December 2009, pp. 29–68;
- Commission Regulation (EU) No 519/2010 of 16 June 2010 adopting the programme of the statistical data and of the metadata for population and housing censuses provided for by Regulation (EC) No 763/2008 of the European Parliament and of the Council. Official Journal of the European Union, L 151, 17 June 2010, pp. 1–13;
- Commission Regulation (EU) No 1151/2010 of 8 December 2010 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission. Official Journal of the European Union, L 324, 9 December 2010, pp. 1–12;
- On the statistical interpretation of the EU regulations on Population and Housing Censuses 2011. Questions & Answers. Luxembourg, 5 August 2013.

The Official Statistics Act passed on 10 June 2010 is the legal basis for censuses in Estonia. Pursuant to the Official Statistics Act, participation in the census is mandatory for all persons permanently or temporarily staying in Estonia. Persons permanently residing in Estonia include those who are temporarily (for up to 12 months) in a foreign country; and public servants serving in foreign missions of the Republic of Estonia and their family members. There was only one optional question and it concerned religion (respondents could choose the option "Do not wish to answer"). All other questions were mandatory for persons enumerated in Estonia, irrespective of whether the question was a core question or not.

Choice of census methodology

One could assume that the arrival of the Internet and computers has made it much easier to enumerate people, compared to the time a century ago when enumerators went from door to door and recorded everything on paper. However, this assumption is far from being true. Firstly, the requirements change with each census – the census questionnaire becomes longer, quality standards become stricter. Secondly, enumeration is more difficult due to increased mobility, advanced self-awareness and desire for privacy. More and more people do not want to be enumerated, as they perceive it as a violation of their privacy.

In Estonia, the preparation period for the 2011 Population and Housing Census (PHC 2011) lasted about five years. The first fundamental issue was the choice of methodology. Already after the previous census in 2000 (PHC 2000), there were articles in the media saying that it was the last traditional census and that all future censuses in Estonia would be based on register data – just like in Finland, Sweden and some other countries with advanced register systems. After all, Estonia already had the Population Register, the register of people insured by the Health Insurance Fund, Pension Register, Register of Taxable Persons; and other state registers were available, planned or being created. However, the data collected from ministries in 2006 showed that all the required information was still not available from Estonian registers and databases.

The question of methodology returned to the agenda at the end of 2008 with recession looming. The government thought that a register-based census would be significantly cheaper than an ordinary census, but they did not consider the work and expenses involved (in order to make the registers suitable for census-taking). Statistics Estonia analysed Estonian registries – Population Register (RR), Estonian Education Information System (EHIS), State Register of Construction Works (EHR), Register of Taxable Persons (maintained by the Estonian Tax and Customs Board), State Pension Insurance Register (PKR), Health Insurance Database (KIRST) – and also considered the Medical Birth Register (MSR) analysed by the Estonian Institute for Population Studies. Based on this analysis, the following conclusions were drawn (as at 2008):

- In all the registers containing personal data, the records are linked to a personal identification code. This ensures the good interoperability of different registers.

- A lot of work was done to develop the address standard: address data system (ADS). However, it was not ready for full implementation before the census.
- The Population Register (a central register where data quality is remarkably good) has one major shortcoming – the address specified in the Population Register differs from the actual place of residence in case of one fifth of permanent residents in Estonia. Therefore, use of the address data in the Population Register would cause errors in population distribution and migration data; it would also distort the composition of households and families determined on the basis of dwellings.
- Education data are only available for about 20% of younger residents, and that information does not fully reflect the education acquired abroad.
- None of the registers contain information about occupation.
- There is no register data about the location of workplace (in case the person does not work at the head office).
- Migration data in the Population Register date back to 2000; there are no records of earlier migration (including external migration) in registers.

The Estonian Academy of Sciences also emphasised the importance of data quality in the census and referred to some problems with Estonian registers. Thus, the Government of the Republic decided that a combined method shall be used to conduct the 2011 Population and Housing Census.

Only 6 out of 51 UNECE countries (incl. all EU Member States) managed to organise the 2010 population and housing census based on register data alone. These countries were Austria, Denmark, Finland Norway, Slovenia and Sweden. Three countries – Belgium, the Netherlands and Iceland – obtained the necessary data by combining register data with previous surveys. The rest of the countries had to carry out an interview census. In six countries – Israel, Germany, Poland, Spain, Switzerland and Turkey – the register data were supplemented with additional surveys conducted to study specific census characteristics. The combined method (featuring the traditional interview census) was used in Estonia and also in Lithuania, Liechtenstein and Latvia. Nine countries used register data only to establish and validate the frame. 22 countries (incl. the United States of America and France who had a significantly different census methodology) conducted the census as a traditional interview census without using any register data (Census methodology).

It should be mentioned that all countries using either a register-based (9 countries) or combined (10 countries) methodology employed quite a large number of registers – 7.4 and 4.9 registers on average, respectively. The population register was used by all. All the nine countries conducting a register-based census used a building register and an employment register.

What does the combined method mean in Estonia? Combining data sources

The combined method offers various options. Firstly, it means **combining data sources**. In principle, two types of data sources were combined: information received directly from the enumerated persons and previously recorded information available in registers (incl. the database of the 2000 census).

The preparations for the 2011 Population and Housing Census (PHC 2011) included an assessment of registers, resulting in the decision to use the following registers in different stages of the census and for different purposes (some of the registers were in the process of being established at that time):

- Population Register (RR)
- Estonian Education Information System (EHIS)
- Database of 2000 census
- State Register of Construction Works (EHR)
- National Defence Obligation Register (KVKR)
- Health Insurance Database (KIRST)
- Register of Social Services and Benefits (STAR)
- Social Insurance Board's information system (SKAIS)
- Register of Taxable Persons maintained by the Estonian Tax and Customs (EMTA)
- Medical Birth Register (MSR)
- Register of Prisoners (VANGIS)
- Land Register (KR)

These registers were supplemented by a few lists (e.g. from the Ministry of Foreign Affairs, list of homeless persons, etc.). A crucial step for the census was the development of the uniform address data system (ADS) by the Estonian Land Board in 2007, although it had not been fully implemented by the time of PHC 2011.

Register data were planned to be used for four purposes, above all:

- determination of enumeration areas and preparation of enumeration lists (in the preparation period);
- addition of information to the census database (without conducting a survey);
- pre-filling of questionnaires (subject to verification);
- supplementing census results in case of missing data; and assessment of coverage.

The preparation of enumeration lists based on the population register is a traditional method which has been used before (the so-called address board information was used in pre-war Estonia; card registers containing detailed addresses also existed in the Soviet times). The Population Register and the State Register of Construction Works were the main registers used for this purpose.

Information about studies and educational institutions was transferred directly from the register (EHIS) to the census database. The birth database (date of birth of first child) of Statistics Estonia and PHC 2000 data were also used.

In Estonia, pre-filling has worked well for income tax returns, where the majority of data is pre-filled and the person submitting the form only has to confirm that the data are correct or modify the data if necessary. Such pre-filled forms, subject to verification, are also used in case of surveys where respondents are interviewed repeatedly. There were several census characteristics the values of which were already available in registers (for most people), which allowed pre-filling. In case of the personal questionnaire, these characteristics were:

- legal marital status (data source: Population Register – RR),
- citizenship (RR),
- mother tongue (PHC 2000, RR),
- country of birth (PHC 2000, RR),
- mother's country of birth (PHC 2000, RR),
- father's country of birth (PHC 2000, RR),
- grandparents' country of birth (PHC 2000, RR),
- place of permanent residence at the census moment in 2000, settlement level (PHC 2000).

In case of the e-census, the respondent was able to view his/her personal information and the information on his/her children aged under 15, provided that the parental connection was established in the Population Register.

However, due to data protection requirements, pre-filling could not be used as much as expected, because the pre-filled data could only be shown to the person concerned. If a person filled in the questionnaire for another household member, he/she was not allowed to see the pre-filled data due to confidentiality requirements. For technical reasons, pre-filling was not available for people living at an address that differed from their registered place of residence.

Registers have an important role in data review after enumeration. In this stage, registers are used for two purposes. The first is filling in the gaps in collected data, i.e. replacing missing items with register data. There are various methods for dealing with data gaps, generally known as imputation. It is possible to use the collected census data and calculate the probable value of a missing item using mathematical formulas. Understandably, data quality is usually better if there are more directly measured characteristics and few imputed values. Use of register data for filling data gaps is not considered imputation if the census uses a combined methodology (because the use of registers for supplementing census data is specified in the methodology). During post-enumeration data revision, the same registers that were used for pre-filling were used for replacing missing items, but to a greater extent – this time it was not limited to the characteristics and registers listed above.

Another post-enumeration task that makes use of registers is the assessment of the general population, i.e. taking account of non-enumerated persons. Such persons were added to the number of enumerated persons (so-called census population), based on information that they were included in registers. This general population was used to assess the quality of the census, but the added persons (objects) were not included in the census results (Tiit 2012).

Besides register data, data from surveys (Estonian Labour Force Survey) were also used to determine the length of the working week of employed persons (a voluntary characteristic).

Combined census-taking methodology

After it had been decided that the census would use a combined method, it was necessary to decide and approve the data collection method. There are two main options: self-enumeration (via an online or mail survey) or interviewing (using printed or electronic questionnaires or telephone interviewing or other means). In Estonia it was decided to use a combination of the following methods:

- self-enumeration over the Internet (CAWI – Computer Assisted Web Interview), which was called the e-census;
- enumeration of residents using laptops (CAPI – Computer Assisted Personal Interview).

There were two backup options: Paper Assisted Personal Interviewing (PAPI), mostly applicable in case of IT problems; and Computer Assisted Telephone Interviewing (CATI) in case of hard-to-reach locations (such as small islands, very remote rural households). In fact, telephone interviewing was used during the few final days of the census period at the end of March, when people called and said that they had not been enumerated and there was no time to visit them. The telephone interviews used a slightly shorter version of the questionnaire which excluded some more detailed questions.

Table 1. Distribution of data collected in UNECE countries during the 2010 census round by data collection method
(percentages)

Country	PAPI (interview)	PAPI (self-enumeration)	CATI	CAPI	CAWI	Registers	Surveys
Albania	100						
Armenia	100						
Austria						100	
Azerbaijan	100						
Belarus	100						
Belgium						98	2
Bosnia and Herzegovina	100						
Bulgaria	59				41		
Canada	12	30	2		55		
Croatia	100						
Cyprus				100			
Czech Republic		73			27		
Denmark						100	
Estonia				32	67	1	
Finland						100	
France	1	99					
Georgia	100						
Germany	15	30			5	50	
Greece	100						
Hungary	65	16			19		
Iceland						97	3
Ireland		100					
Israel		10	20	60		10	
Italy		67			33		
Kazakhstan	100						
Kyrgyzstan	100						
Latvia				59	32	9	
Liechtenstein		65			25	10	
Lithuania	61				34	5	
Luxembourg		100					
Malta	99		1				
Moldova	100						
Montenegro	100						
Netherlands						95	5
Norway						100	
Poland			3	22	12	63	
Portugal		50			50		
Romania	100						
Russia	100						
Serbia	100						
Slovakia		93			7		
Slovenia						100	
Spain		48		10	37		
Sweden						100	
Switzerland		2			1	97	
Tajikistan	100						
Turkey	32			48		20	
Ukraine	100						
United Kingdom		84			16		
United States of America	33	63					

Source: UNECE

In case of censuses, the most traditional method of data collection is paper-assisted interview (PAPI). This method has been used for most censuses conducted in Estonia. It was also used in the Soviet Union and has been the chosen method in Eastern Bloc countries.

In case of self-enumeration, the respondents fill in the questionnaires by themselves without the enumerator's presence (as a rule). This method covers mail surveys and online questionnaires (CAWI). Naturally, self-enumeration involves a greater risk of receiving erroneous or even intentionally distorted responses, but self-enumeration has been used in Western Europe for several decades and there have been no major problems (based on survey reports). Self-enumeration was tried out in Estonia in 1922 and 1934, when the enumerator distributed the questionnaires to respondents and returned later to collect the completed questionnaires. The main risk of self-enumeration is false response, given either intentionally or unintentionally. This means that utmost attention has to be paid to the wording and understandability of questions, help texts and support systems. In order to reduce the probability of random errors and intentional fooling, the online questionnaires were equipped with systematic logic checks.

Obviously, it is not yet possible to organise a census using only online enumeration. Additional interviewing is required for those who do not have the possibility or do not want to complete the census online. In Estonia, online self-enumeration was combined with laptop-assisted interviewing, because Statistics Estonia already had the experience of conducting surveys using laptops. Use of laptops instead of printed questionnaires was also considered more economically efficient. Table 1 outlines the use of various data collection methods in UNECE countries.

A significant advantage of both online surveys and computer-assisted interviews is the fact that the data are collected in a database without a separate data entry phase (even during the previous census in Estonia, data entry had to be performed separately) and that logic checks can be used.

International comparison revealed that only Estonia and Canada collected the majority of census data was collected via online self-enumeration. In Portugal, respondents could choose between two self-enumeration methods – online or on paper – and both options were used almost equally. Online enumeration was used, to a smaller or greater extent, in 13 other countries. Seven countries used computers and four countries used telephone interviewing for data collection. However, even in this census round, the most popular method was use of paper questionnaires – these were used in 37 countries, whereas fifteen countries asked the respondents to complete the questionnaires themselves (in some cases, this did not apply to all respondents). 26 countries managed with only one data collection method: six countries used only registers, one country (Cyprus) used only computer-assisted interviewing, and 18 countries used paper questionnaires (self-enumeration was used only in Luxembourg). Thus, developed countries mostly used a combined method of census-taking in this census round.

Choice of census moment

The census moment is traditionally the critical moment of the census, i.e. the day, hour and minute to which all events and data refer. To put it simply, all people born before the critical moment are subject to enumeration, but persons born on the reference date but after the census moment are not included. Similarly, the census covers for all persons who were alive at the critical moment but died after that moment (perhaps even on the same day). Place of residence, family structure and other data are also registered as at the critical moment. For instance, if two people got married on the census reference date but after the critical moment, they are not considered as married for the census purposes.

In the European Union, 2011 was the year chosen for the conduct of the 2010 census round – this means that all EU Member States had to organise the census so that the census moment would be within the year 2011. Countries were free to choose the specific date.

In Estonia, the choice of the census moment was subject to very thorough consideration. The census moment (critical moment) of the 2011 Population and Housing Census (PHC 2011) was set to midnight 00:00 on 31 December 2011. Thus, the last day of 2011 was chosen (which happened to be a Saturday). This was also the first day of the census following the critical moment, and the children born on that day were not subject to enumeration.

The choice of this date for the census was supported by several factors.

- Most of the previous censuses – eight out of ten – had taken place near the turn of the year, either in December (3 times) or in January (5 times). Only two censuses in Estonia took place in early spring (March). A significant advantage of winter-time censuses is the fact that the population is less mobile than in the summer. Another aspect is related to the enumeration of the homeless – they are also easier to enumerate in the winter when they gather at shelters.
- The use of census data in population statistics is easier and the results are more accurate if the census reference date falls close to the turn of the year.

- The development of software for data collection, the preparation of questionnaires and software testing took more time than expected. These tasks were completed only by the end of 2009, which postponed the pilot census (conducted at the end of 2009) and the census moment (set to the absolutely last permitted date).

Census organisers

A new unit was established in the Population and Social Statistics Department of Statistics Estonia: the Population and Housing Census Service with a dozen employees. However, many more employees of Statistics Estonia were involved in the conduct of the census (staff from the Data Collection, IT, Methodology, Marketing and Dissemination Departments etc.).

To ensure the proper scientific foundation of any decisions, the PHC Scientific Council was established. It consisted of 12 members (specialists from universities and research institutions). The Council convened two to four times a year, as necessary, and held occasional e-meetings. The Scientific Council only had an advisory role.

Decisions regarding work organisation were made at the weekly meetings of the PHC Service and the PHC Project Manager. Any major decisions were discussed at the meetings of the PHC project team, attended by all executives in charge of any PHC-related functions at Statistics Estonia. During the active census period, these meetings were held on an almost weekly basis. The next-level governing body was the PHC board, composed of all heads of department at Statistics Estonia and representatives of the Ministry of Finance. The highest PHC authority was the Census Committee of the Government of the Republic, consisting of the chancellors of the ministries involved in the census.

Census information system

A new information system was developed for the census – VVIS (Survey Fieldwork Information System), with the following general lines of action:

- (technical) preparation of questionnaires, incl. the addition of logic checks and help texts;
- collection of census results and storage of the results in a suitable format (compilation of the preliminary database);
- monitoring, checking and documentation of the activities of enumerators, based on the enumeration lists of enumeration areas;
- providing a channel of communication between the different links/levels of the census team;
- ensuring data security.

For security purposes, the system prevented unauthorised access to the data and even the members of the census team (including those engaged in data processing and review) could usually only access a certain part of the data (e.g. only addresses), but never the whole questionnaire of a single person.

How many people would complete the census online?

This was a question that needed at least a rough estimate before completing the final design of the census, because the answer determined the distribution of resources – the number of enumerators required (incl. equipment for enumerators etc.) and the necessary capacity of communication channels during the e-census. Estimates varied a great deal. An external expert assessment (survey company Resta) projected only a 5% participation rate in the e-census. Linnar Viik, an IT-expert and a member of the PHC Scientific Council, proposed the highest response rate – 75%. His optimism was shared by Ene Ergma, the President of the Riigikogu, who believed that that e-census participation rate in Estonia would be at least 50% (she shared this opinion while speaking at the census-related conference held by the Estonian Statistical Society). The experience of other countries was not very helpful in the projection of the e-census participation rate, as there was not much international experience to rely on. A few European countries and Canada (where censuses are organised every five years) had used online enumeration to a small extent in their previous census, but it was unclear whether these results could be comparable to the circumstances in Estonia.

After the pilot census, some model-based calculations were made – these indicated a possible participation rate of 27%, provided that all external conditions remained the same. The PHC team preferred a conservative estimate, choosing 25% as the projected rate of participation in the e-census. However, their recommendation was to use a higher participation rate (40%) when planning the capacity of communication channels. The real e-census participation rate was quite different from the estimates (see Table 1, p. 8).

PHC 2011 QUESTIONNAIRES

What is the purpose of census questionnaires?

The contemporary census represents a comprehensive survey, in the course of which a great many demographic indicators are ascertained. In order to get such results for all the indicators that would be nationally and internationally comparable and help to identify developments and changes, a questionnaire is prepared to get data for the indicators under study, and in most cases, multiple-choice questions are also given.

The population and housing census focuses on three main objects – persons, dwellings and households – and there are also, correspondingly, three questionnaires, but considering that in most cases each dwelling has one household, individual dwelling and household questionnaires can be combined into a single dwelling and household questionnaire (which was also the case for PHC 2000). Usually, uninhabited but inhabitable dwellings are also subject to enumeration.

Besides the main objects, some output tables include derived objects, such as residential buildings. Individual objects, for which data are submitted, include also families, which usually constitute a subset of households and can be derived based on households.

The number and wording of the questions in the questionnaire and the list of multiple-choice questions determine the content of each census. Estonia has undergone censuses with very varied capacity. The scarcest was the census of 1941, when census questionnaires were replaced by tables with only seven characteristics for each respondent – place of residence, sex, age, ethnicity, religion, area of activity and employment. Relatively few questions were used in the censuses performed during the Soviet times (from 1959 to 1989). At first, there were no housing-related questions at all, later on they were asked selectively. The censuses organised by the Republic of Estonia, on the other hand, were much more detailed, especially in terms of living conditions. The most comprehensive one was the census of 2000, which, for the first time, included questions about the parents' place of birth, the average number of weekly working hours and the former place of residence. In terms of the volume of information received, this census was surpassed only by PHC 2011.

Preparation of census questionnaire

Preparing the questionnaire is usually the duty of the PHC organiser – which is typically the statistical office of the particular country – and in most cases expert representatives of both the consumers and researchers-statisticians are also involved. When preparing the census questionnaire, one has to consider that adding each new question will make the already expensive census considerably more expensive. Additionally, an excessively long questionnaire may aggravate the response quality (due to the respondent's boredom or fatigue). On the other hand, the information gained in the course of the census is especially valuable because of its complexity – the examined characteristics can be linked to very extensive background information and, in most cases, it also allows following further developments. In addition, the census creates the opportunity to analyse small population groups, on which there are no other ways to get the information. When choosing the census questions, inevitable disputes arise between the representatives of consumers and organisers, which are the more active, the more open the society and the better the information exchange.

The questionnaire of the first census held on the territory of Estonia was compiled by Professor Schirring from the University of Tartu, taking into account the recommendations of the International Statistics Institute (ISI). The members of the committee preparing the first census in the Republic of Estonia in 1922 examined the census questionnaires and experience of ten European countries. The questionnaire was submitted for discussion to the newly established Statistics Council. In the re-independent Estonia, census preparation has involved expert committees, who voice their opinions about the suitability and necessity of questions included in the questionnaire, and such opinions have been taken into account when compiling the questionnaires.

The census questions were prepared based on three facts and the questions can be divided accordingly:

- International recommendations and requirements, including questions concerning mandatory output characteristics (the so-called core topics described by the United Nations Economic Commission for Europe (UNECE)), in which case all EU Member States are required to submit data in a pre-determined format (scale). The Member States can decide on how the question is asked to get the required information (or how to acquire such information from other sources). For that purpose, such questions required for core topics are referred to as questions imposed by the European Union.
- Questions concerning previous censuses carried out in the country and options for creating comparable time series are referred to as questions traditionally included in Estonia.
- Questions concerning the requests made by stakeholders, including researchers, based on actual survey topics, are referred to as questions that have been newly added in Estonia.

Core questions imposed by the European Union generally represent the most important census questions always asked during censuses. Quite a large share of the relevant information is also stored in registers in Estonia, which increases the possibility of organising the next census based on register data. The concerned topics were declared mandatory for the EU Member States by Regulation (EC) No 763/2008 of the European Parliament and of the Council.

Each census round adds questions about new issues on the international scale, which have become of interest due to the development of the global population. In this round, there were two new fundamental topics emphasised in Europe, which required supplementing the questionnaire. These were migration and the experience of living abroad, and employment (unemployed respondents were asked about previous employment, economic activity, and employment status).

As for **questions traditionally asked in Estonia**, the more important ones include ethnicity and mother tongue. Similar to the previous census, they included a question about the parents' country of birth to describe the parentage structure of the population. In view of Estonian censuses, the question about religion may also be considered a traditional question, although it was excluded from Soviet censuses. As the fertility rate has always been a significant issue in Estonia, all adult women are asked about the number of children they have given birth to. In most censuses (except in the Soviet period), one or more questions were also asked about health, more specifically, about the presence of a disability. While in earlier times the question concerned a particular disability (blindness, dementia), the recent censuses included questions about the presence of a disability and about coping in general. Estonian censuses have also asked questions about the main source of subsistence (in 2000, two sources could be named). The majority of censuses conducted in Estonia have also contained a question about studies (at the time of the census), and in some cases foreign language skills have also been (selectively) asked about.

The questionnaire of the 2011 Census in Estonia also contained entirely **new questions**. The parentage section was supplemented by a question about the grandparents' country of birth (which is rather unique on the international scale). Considering the constantly increasing mobility of people, the respondents were asked about their secondary place of residence used due to studies, work or other reasons. Initially, a separate section was planned for the household's small-scale agricultural production for own subsistence (not included in previous censuses), but after the pilot census it was discarded and replaced by a single question in the household questionnaire. New questions were also added to the list of core questions, e.g. which country was the respondent's previous place of residence, the unemployed were also asked for their last year of employment. The question about the type of heating was also more detailed than before.

An in-depth discussion followed the proposal of the local stakeholders to add a question about the ability to speak a dialect. As the issue of dialects has recently been topical in Estonia, the census team decided to add the question to the questionnaire.

One of the most significant last-minute supplements to the questionnaire was the question about close relatives who had left to live abroad. Emigration is not usually covered by censuses, because the questions can only be asked from the residents of the country concerned, and not from the persons who have left. However, the information gained from households about close relatives who have left for abroad allows specifying the number of people who have left to live abroad. All the added questions were approved by the PHC Scientific Council.

Figure 1. Distribution of questions in the personal questionnaire of PHC 2011
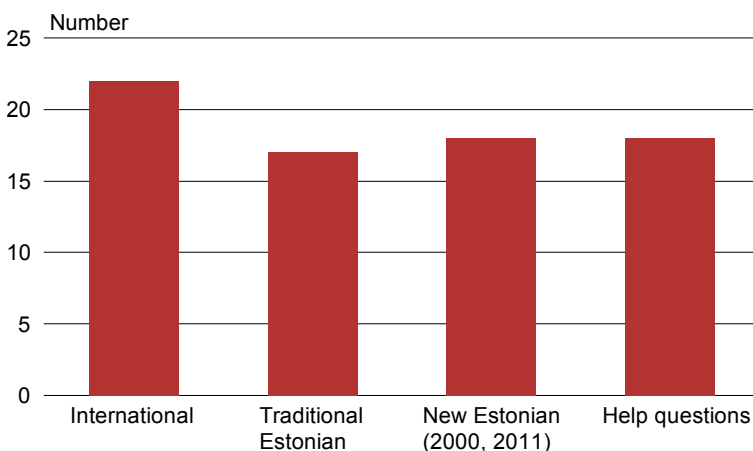
Rejected proposals. As always, a few questions proposed by the stakeholders had to be rejected this time as well. Another question that was proposed along with the one regarding dialect concerned the ethnic group, but as this concept is not well-known, it was left out of the questionnaire. The Estonian Society for the Protection of Animals wished to add

questions about pets – and not just in terms of the existence and number of dogs and cats, but also with regard to pedigree and chip-related data. Dissatisfaction was also expressed with the wording or multiple-choice answers of several other questions (emphasising the native religion in relation to the question on religion, ethnic nationality was suggested to be considered as citizenship, etc.).

How many questionnaires were created for the census?

A great many questionnaires had to be compiled for PHC 2011. As there were three basic enumeration objects – person, dwelling and household – there were also three main questionnaires, although the household and dwelling questionnaires (in hard copy) were combined into one, i.e. their separation was more of a virtual one.

All the above-mentioned questionnaires were prepared in three versions: as an online or CAWI-questionnaire, a laptop-assisted interview or CAPI-questionnaire, and on paper, i.e. as a PAPI-questionnaire. It was extremely important to have the same wording and order of questions in all three versions, so that the responses would not depend on the method used for filling in the questionnaire. Inevitably, there still occurred some discrepancies. The online questionnaire started with the identification of the respondent (preceded by a few questions about the dwelling). Interview questionnaires, however, started with the list of household members. The self-enumeration questionnaire required more help texts and explanations, the CAPI-questionnaire used less explanations. In the CAPI-questionnaire, the respondent could refuse to answer single questions, in which case the enumerator indicated the response "does not know" or "refuses". In the case of the online questionnaire, that option was more limited, but even then it was possible to leave a question unanswered by ignoring the "soft" checks used for drawing attention to the questions missed. Logic checks could not be applied in the case of paper questionnaires.

In addition to the ordinary personal questionnaire, there were shorter (virtual) personal questionnaires, containing a few questions about a temporary resident and a person who had left Estonia. These questionnaires were usually filled in not by the respondent himself (it was generally impossible), but the representative of the household instead, i.e. these constituted a part of the household questionnaire.

All questionnaires were translated into Russian and English, the paper questionnaire was also available in Finnish. Thus, each respondent was able to pick the most suitable language. This continued an old tradition – the first censuses carried out on the Estonian soil (in the 19th century) already had questionnaires available in four languages: Estonian, German, Russian, and Swedish.

Great attention was paid to the wording of the questions. In summer 2008, 24 questions were tested by using the cognitive interview method. The test also took into account the amendment proposals made at the approval round of the questionnaire in spring 2009, and the experience gained from the pilot census of 2009 and the mini-pilot census of 2011. Questionnaire compilers aimed for maximum accuracy, especially with regard to the definitions. This caused frequent protests at the first reading – the wording felt clumsy, too long and difficult to read. The wording of the questions and answers was polished and clarified during the entire preparation period that lasted for several years. The questionnaires were particularly problematic in the Russian language, with even the key concept "household" being perceived as difficult to understand by the respondents.

In the case of electronic questionnaires (both online and for the laptop-assisted interview), the inner logic of the questionnaire was of utmost importance, so that the respondent was not presented with any unnecessary questions, i.e. questions that the respondent did not have to answer (e.g. men were not asked how many children they have given birth to). In the case of paper questionnaires, the enumerator had to skip such unnecessary questions.

Help texts constitute an integral part of the questionnaires, and they are particularly important in the case of self-enumeration. In this census, the online questionnaires utilised help texts on various levels – in addition to those constantly visible on the screen (in limited number and volume), more detailed help texts were available for almost every question.

Several questions in the online questionnaires and on the enumerator's laptop allowed answering in three principal stages: in the first stage, the respondent could choose from only a few (3-7) predetermined multiple-choice answers. The majority of the respondents were able to answer the question using that option. If there was no suitable choice among the pre-set answers, the next step involved using a classification that opened in a drop-down menu (in the case of an interview, this was done by the enumerator). If a suitable answer was still not found, the respondent could use a comment field to enter the answer (filled in either by the respondent or by the enumerator). This option could be used for answering questions about citizenship, ethnicity, mother tongue and several other characteristics.

In addition to that, the electronic questionnaires contained logic checks as well. To some extent, logic checks have been used in all censuses – they have been applied to verify (even manually) the age relations between parents and children and other obvious connections. In earlier censuses, after discovering a mistake, the doubtful information was usually referred back to the respondent for further clarification. Electronically filled questionnaires allow using logic checks at the

time of filling in the questionnaire. There are two types of checks. A "soft" check informs the respondent of an unusual or unlikely but still possible situation (Your age is 23. Do you have a doctoral degree? The spouses' age difference is 35 years. Is this correct?). Filling in the questionnaire can be continued after responding to a "soft" check, either by confirming the current situation or correcting the mistake. A "strict" check notifies the respondent that the situation stated in the questionnaire is impossible (the father is younger than the son), and the questionnaire cannot be filled in until the mistake is corrected. Thus, electronic checks also excluded most of the instances of fooling around or joking. A skipped question in the online questionnaire generally activated a "soft" check, which reminded the respondent of the skipped question, but still permitted to continue with filling in the questionnaire.

In the case of questions that were mandatory in a technical sense, the respondent was not allowed to continue filling in the questionnaire if any of the questions were unanswered. All that ensured the good quality of census data.

The total number of checks used in the PHC questionnaires was remarkably high: CAPI used approximately 280 unique checks (124 of them were household matrix checks), and CAWI used approximately 310 unique checks (122 of them were household matrix checks). Some checks were activated in several places (e.g. address checks, checks for the type of the dwelling regarding the main dwelling and empty dwelling of the household, checks for the personal identification code and date of birth), resulting in the higher total count of checks used in the questionnaire – an estimated 342 checks in CAWI and 345 in CAPI. The checks were mostly the same in CAWI and CAPI, but CAWI had more reminder checks.

The Annex contains the paper versions of both the personal questionnaire and the household-dwelling questionnaire in Estonian. Although hard copies were rarely used in the census (only in the case of the so-called emergency situations) and their structure is in a way the most primitive, they were mostly presented for introduction purposes, as it was technically the simplest way and required no additional devices.

CONTENTS OF THE PHC 2011 QUESTIONNAIRES. QUESTIONS IN THE QUESTIONNAIRES

The classification and order of questions follow the headings described at the Conference of European Statisticians in 2006 and the indicated core and voluntary characteristics (UNECE 2006). The questions presented in the questionnaire fall into eleven groups, with each of them distinguishing between core and voluntary questions. The definitions used in the text below originate from the explanations added to the information about PHC 2011 in the Statistical Database (under Definitions).

Enumerated persons and their permanent place of residence

The key issue in finding the size of the population based on census data is who to enumerate. In principle, there are two options – either to enumerate the **de facto population**, i.e. all the persons present in the country at the census moment (even if they are just passing through), or the **permanent population**, i.e. all the persons residing in the country permanently. While early censuses focused on the de facto population, recent censuses prefer the permanent population. The Census 2000 considered both types of population, but the main output tables were still based on the permanent population. In the 2011 Census, enumeration covered only the permanent population in Estonia.

The permanent residents are determined according to their **place of residence**, the definition of which is rather complicated. The place of residence is determined by an address (in Estonia, it contains 8 components based on the standard set by the Address Data System (ADS), and spatial coordinates). The place of residence is a core characteristic that has to be determined for every enumerated person.

The place of residence (permanent/usual place of residence) is usually the region or settlement, where a person spends most of his/her daily rest and sleep time and where the person has been living continuously for at least 12 months before the census moment, or where he/she came to live before the census moment, and intends to stay for at least one year. The definition is complicated by the fact that, these days, people may have several places of residence (incl. in different countries), while only one of them can be indicated as the permanent place of residence for census purposes.

- Persons who lived away from home due to work for more than 12 months but spent most of their days off with their households were considered to be residents at the address of their household.
- Pupils of general education schools and secondary vocational education institutions who resided outside home during their studies were generally considered to be residents at the address of their household.
- Students of institutions of higher education and post-secondary vocational education residing outside their former homes during their studies were generally considered to have a place of residence at the address at which they lived during the studies, provided that they lived there for the majority of the study period. If a student had his/her own (new) family, then his/her place of residence was the residence of his/her family.
- Households (including one-member households) who regularly lived at several addresses during the year were considered to have the place of residence at the address where they spent the majority of the year.
- Persons residing in an institution (care home, custodial institution, etc.) were considered to be residents of that institution if they had been living in the institution for at least 12 months or more or would be staying there for more than a year.
- Persons in compulsory military service were generally considered to have their place of residence at the address where they lived before military service.
- The military and navy staff of the Republic of Estonia, diplomatic staff and their family members who live with them who were in a foreign country during the Census were considered to have their place of residence in Estonia.
- Persons without a place of permanent residence (incl. the homeless) were considered to have their place of residence at the place (dwelling, settlement) where they were at the census moment.

The census covered all permanent residents of Estonia, i.e. persons whose place of residence was in Estonia at the census moment. Thus, according to international rules, the permanent residents of Estonia also included a certain amount of commuters – persons who work in another country but spend most of their free time with their family. Meanwhile, persons who tried enumerating themselves online while staying in a foreign country and indicated that their permanent place of residence was in a foreign country were not considered to be permanent residents of Estonia (there were more than 4,300 such persons, i.e. 0.3% of the total enumerated population). The basic principle of the census establishes that the word of the person enumerated is considered to be true, and it is only verified in the case of an obvious contradiction.

All the enumerated permanent residents had to indicate the address of their permanent place of residence according to the ADS-standard, which also included the spatial coordinates. Classifications (EHAK – Classification of Estonian Administrative Units and Settlements) were used to facilitate the insertion/selection of different address levels. For various reasons (registering one's address for the first time, limited knowledge of the ADS, technical issues, being unused to marking spatial coordinates), the address information received required rather extensive clarification and specification, which took more time than initially anticipated. All in all, it can still be confirmed that the information received on permanent places of residence was complete. Persons without an exact dwelling address, incl. the homeless, were considered to live in the middle of the settlement (in a fictitious building, a so-called settlement centroid). Such dwellings totalled 4,086, i.e. 0.6% of all dwellings.

Based on the number of enumerated persons, the total population (1,294,455) was found. This is the so-called census population, which may contain mistakes made in the course of the census (under-coverage and over-coverage), but is still used as a basis for the compilation of all output tables.

In addition to information about the permanent residents, data was also collected about certain groups of the population not included among the permanent residents of Estonia. These were temporary residents whose permanent place of residence was abroad. A temporary resident was a person who lived (intended to live) in a given place of residence for 3–12 months. Another population group enumerated but not considered to be Estonian permanent residents were persons who, according to their close relatives, had moved abroad after the year 2000 and stayed there (the so-called leavers).

The enumerated temporary residents may also have included Estonian permanent residents. For example, parents might have indicated their child, who was a university student, as a temporary resident at their place of residence. Temporary residents also included Estonian permanent residents who were not in their permanent place of residence most of the time during the census. Such persons were generally also enumerated in their permanent place of residence and their permanent place of residence in Estonia was clarified in the course of data processing.

During PHC 2011, the respondents were, for the first time, asked about the presence and location of a possible **secondary place of residence** (only if it was not located in the same city/rural municipality as the permanent place of residence). The secondary place of residence is not a core characteristic.

Geographical characteristics

The definition of permanent residents is also applicable to smaller geographical areas (county, rural municipality, city, and village). They are all characterised by population or the number of permanent residents living in the particular region, the further examination of which uses the so-called geographical characteristics.

Traditionally, population is divided into rural and urban population. Today, urban sprawl and other migration processes have resulted in rather vague borders between urban and rural areas. Therefore, the traditional urban/rural distinction is deemed voluntary and has been replaced (or supplemented) by the concept of **locality**, which denotes a region that has a certain settlement but is distinct from its surroundings; it could be a city, town or village, or merged (conjoined) settlements. A city can contain several localities if they are clearly divided, for example, by a river (in Pärnu) or a park zone (Pirita city district in Tallinn). A locality is characterised by the population number determined according to the standard scale, of which, the following section is used in Estonia:

- 200,000–499,999
- 100,000–199,999
- 50,000–99,999
- 20,000–49,999
- 10,000–19,999
- 5,000–9,999
- 2,000–4,999
- 1,000–1,999
- 500–999
- 200–499
- < 200

Locality is a derived characteristic, the determination of which requires fixed spatial coordinates of the place of residence.

In addition to the locality of the permanent place of residence, the group of geographical characteristics also includes the location of a person's job and educational institution. **Location of job** (country of job and city/rural municipality of job) is a core characteristic and refers to the actual place of work (local activity unit), which is generally not available in Estonian

registers. The country of work was well-indicated (the information was missing only in the case of 0.1% of the respondents), whereas the exact location of the workplace in Estonia was rather poorly indicated – 6% of the respondents left it unanswered. The reasons for that include both indisposition to respond and (more frequently) unawareness, especially if the questionnaire was completed by another person. Classifications were used for all geographic characteristics. In addition to the currently valid administrative division, the respondent could use former ones, and if a suitable designation was not included in the classification, then it was possible to enter the information in the comment field. Such free-text answers were encoded during the data processing phase.

Location of educational institution is not a core characteristic and it was filled in based on the data recorded in the Estonian Education System (EHIS).

Demographic characteristics

The core demographic characteristics include **sex** and **age** (date of birth), which also represent the characteristics used to identify a person. The enumerated persons were asked for their personal identification code, which contains the full information about their gender and date of birth. The personal identification code was used to calculate the person's age, which was then verified by the respondent. If the respondent did not indicate his/her personal identification code (in the online census it was only possible if the person did not enter the census environment by himself/herself), then the respondent (or the person responding on his/her behalf) had to enter the sex and date of birth in the questionnaire or tell it to the enumerator. Sex is included in almost all output tables, and age in most of them. All previous census questionnaires contained a question about sex and age. The date of birth is included in one output table. Due to the suitably chosen census moment, the difference in the information related to age and date of birth contained in the tables of PHC 2011 is insignificant (by one day or 0.3%). Sex and age were used to delimit population groups when completing the census questionnaires.

Another core demographic characteristic is **legal marital status**. To pre-fill this characteristic and to find the missing values for it, the Population Register was used. The share of replaced values was 0.7%, but the legal marital status was still missing in the case of 1.2% of the respondents. As the legal marriage or other registered cohabitation of persons of the same sex is not permitted in Estonia, the legal marriage scale is simple and contains only four categories (never legally married / legally married / divorced / widowed).

De facto marital status is not a core characteristic, but it is still very important in Estonia today. It was not asked as a separate question, but was included as one option in the household relationship matrix, where every household member was asked about the existence of a partner and whether it was a legal spouse or an unregistered cohabitation partner. In the case of de facto marital status (unlike legal marital status), only the current situation is recorded. The partners may also include same-sex partners.

The demographic characteristics that are traditional for Estonia also include the **number of live-born children** (preceded by the ancillary question – have you given birth?) and **age at the birth of the first child**. These data were calculated for women aged 15 and older. 0.9% of the respondents did not answer the question about the number of live-born children (0.8% of the gaps were filled from the PHC 2000 database). Age at the birth of the first child was retrieved from registers, the information was absent in 3.4% of the cases. Considering that fertility has always been an important issue in Estonia, the number of live-born children has been included in all censuses organised by the Estonian state.

Economic characteristics

The census questionnaire contained five core characteristics related to the respondents' economic situation: **labour status, socioeconomic status, employment status, occupation and economic activity (of the workplace)**; plus three characteristics requested by Estonian stakeholders: **main source of subsistence**, **typical length of the working week** and **in the case of unemployed persons their last year of employment.**

The first question for ascertaining the **socioeconomic status** concerned a person's employment during the reference week (asked from persons aged 15 or older). This census used a single reference week – from 19 to 25 December 2011 – for all respondents, and even at least one hour of remunerated work during the reference week counted as employment. The people who had not worked on the reference week were then given a set of three questions, developed by the International Labour Organisation (ILO), to determine their **labour status** (absent from work / seeking work / ready to commence work). These were followed by questions about previous employment and the year of last employment. No registers were used to supplement the labour status data. All in all, labour status remained unknown for 1.3% of the population. Based on labour status, the population is divided into the employed, unemployed, and economically inactive population. The employed and the unemployed constitute the economically active population. The economically inactive population (retired persons, students, etc.) was further specified by additional questions. With regard to this question, the chosen time (working on the Christmas week!) was relatively inconvenient – at this time,

occasional work was probably done less and job searching was not as active either. This resulted in a discrepancy between the results of the census and the Labour Force Survey in terms of the number of employed and unemployed persons. **Socioeconomic status** describes both economically active and inactive persons. Census information was not supplemented by using any other sources for this characteristic either and the answer was missing in approximately 1.3% of the cases. **Employment status** is only used for making a distinction between employees (employee, entrepreneur, etc.). In several output tables, the labour status, socioeconomic status and employment status have been combined into a single characteristic **employment and socioeconomic status.**

Only employed persons and those unemployed persons who had worked before had to answer the questions about their employment status, occupation and the economic activity of the main place of work. The questions concerning occupation and the economic activity of the main place of work were significantly different from output indicators. Respondents had to mark down the full name of their main place of work, and – if necessary – of the subordinate unit (essentially only an ancillary characteristic). In order to determine the particular economic activity of the main place of work, the respondents were asked to describe it (in addition to naming the economic activity). The same applied to occupation – in addition to naming the occupation, the respondents had to describe their work-related duties. If the description of the occupation suggested that the respondent might have subordinates, the number of subordinates had to be specified as well.

Text answers were encoded over the course of data processing and the results were presented in accordance with international classifications (NACE, ISCO-08). Economic activity was published on the level of one-, two- and three-digit classification codes. On the one-digit code level, economic activity remained unknown for 0.8% of the respondents. Occupation was published on the level of a one-, two-, three-, and four-digit classification code. On the one-digit code level, economic activity remained unknown for 1.6% of the respondents.

A similar question layout about economic activity and occupation was also used in the Population and Housing Census 2000; both the experts of Statistics Estonia and several consumers requested classification in as great detail as possible (exceeding the mandatory requirements). The encoding of occupations and economic activities was relatively labour-intensive and most of it was performed manually.

The **main source of subsistence** was asked about from all persons aged 15 or older. Younger persons were generally considered as "maintained by other persons", except for children living in institutions (e.g. children's homes), whose source of subsistence was marked as "maintained by institution". The question of the main source of subsistence has been included in most of the population censuses conducted in Estonia, and the previous time even two main sources of subsistence were asked about. This question remained unanswered for 1.3% of the respondents.

Persons qualified as unemployed were asked about their **previous employment and duration of unemployment**. The question of previous employment remained unanswered only in 0.3% of the cases, but 1.2% of those who had had previous employment, did not report the year of last employment. This question was included in the census questionnaire for the first time in Estonia.

The **average length of the working week** was not asked about in the questionnaire and the data of the Labour Force Survey (the average of 2011) were used instead.

Education characteristics

Only one of the education characteristics is a core characteristic – **educational attainment**, i.e. the highest level of education completed. Internationally, this is measured by using the 1997 version of the International Standard Classification of Education (ISCED):

- primary education (level 1)
- lower secondary education (level 2)
- upper secondary education (level 3)
- post-secondary non-tertiary education (level 4)
- first stage of tertiary education (level 5)
- second stage of tertiary education (level 6)

Acquisition of education reputedly comprises a great many potential combinations of general and vocational/professional education. In order to describe educational attainment on an international scale, while presenting the levels of education in a manner that is comparable to the previous population censuses carried out in Estonia, a questionnaire consisting of six questions was developed. Besides the usual questions about the attainment of general, vocational, and professional education, the respondents were also asked about the preliminary education required for the respective educational institution, as it was necessary for determining the ISCED level. At the experts' request, respondents with no elementary education were also asked about **literacy**. The questions were asked from persons aged 15 or older, the information

concerning those aged 10–14 was gained from the Estonian Education Information System (EHIS). Missing information was supplemented (where possible) with data from EHIS, but the level of education still remained unknown for at least 1.4% of the respondents aged 10 or older.

In addition to that, the census database (and output tables) was supplemented with information about the **level of education being acquired** by students attending the educational institutions located in Estonia at the census moment.

Internal and external migration

In this census round, great attention was paid to migration questions because people are very mobile these days, which, among other things, makes organising a census unprecedentedly complicated. Migration was associated with the following core characteristics:

- country of birth and place of birth;
- citizenship;
- living abroad and year of immigration/returning;
- previous place of residence and time of arrival at current place of residence.

In Estonia, the PHC 2011 questionnaire also included the following questions (characteristics):

- country of previous place of residence;
- usual place of residence at the time of the previous population census in Estonia (31.03.2000);
- father's country of birth and mother's country of birth;
- grandparents' country of birth;
- close relatives who have emigrated abroad in 2000–2011.

The definition of the **place of birth** proved to be somewhat complicated – it refers to a place where the person's mother was a permanent resident at the time of the person's birth. In some cases, this is different from the previously known place of birth (which had even been entered in the passport), which often corresponded to the location of the maternity hospital. The place of birth entered in the Population Register is also inaccurate (at least in some cases), because it does not comply with the definition featuring the place where the person's mother was a permanent resident. Both the data from the Population Register and PHC 2000 (2% and 0.3%, respectively) were used to supplement the information about the country of birth and the place of birth, and the question remained unanswered for 0% and 1.8% of the respondents, respectively. The question regarding the place of birth has also been included in Estonian censuses before. Here, too, classifications were used for all geographic characteristics; a more detailed place of birth was not required for those born in a foreign country.

As for **citizenship**, Estonia stands out for a large number of persons with undetermined citizenship. The Population Register generally contains accurate and up-to-date information about citizenship. Therefore, it was possible to pre-fill the citizenship question based on the Population Register data and use the same source to fill the gaps in the data. Information was specified and supplemented in 2.1% and missing in 0% of the cases. Citizenship was also indicated with the help of the international classification of countries.

The question about **living abroad** was only asked from persons who had not been born abroad. This question remained unanswered by 1.7% of the respondents and the time of arrival was not indicated by 2.1% of the respondents. A somewhat better response quality was observed in the case of the question about the most recent country of residence (1.6% of responses missing).

The **country of the previous place of residence** was covered well (only 1.2% of the responses missing), while there was a serious lack of information regarding the address of the previous place of residence – values were missing in the case of up to 13% the respondents. This is another problem with reference to geographic information, possibly due to its (technical) complexity, but also due to forgetting. The response quality was only slightly better for the question regarding the **time of arrival at the last place of residence** (values missing: 9.7%). No external sources were used to adjust the values of these characteristics.

However, a wealth of information was provided by a characteristic voluntary in Estonia – the **country of the place of residence** and the **place of residence in Estonia at the time of the 2000 Population Census**. These questions were partially pre-filled. The country of the place of residence was not indicated by only 0.5% of the respondents, and the settlement of the place of residence of those who lived in Estonia at that time was not indicated by 0.1% of the respondents. For these characteristics, information was supplemented by using the database of the Census 2000.

Considering the interests of Estonian demographers, questions regarding the respondents' native origin were also added to the questionnaire. As we follow the common division of characteristics applicable in Europe, according to which all birthplaces are classified as migration characteristics, the characteristics regarding native origin are shown under this and not the next heading. The question about the **parents' country of birth** was also included in the previous census, this time a question about the **grandparents' country of birth** was added. After supplementing the information about

these characteristics based on the Population Register and the PHC database (to an extent of 1.3% on average), the mother's country of birth remained unknown for 0.6%, the father's country of birth for 1.3%, and the grandparents' country of birth for 4.6% of the respondents. Based on the grandparents' country of birth, that of the parents' and the person's own country of birth, the Estonian population was classified into the native population, and the immigrant population of the first, second and third generation.

Information was also collected about the **persons having emigrated abroad in 2000–2011**. The representativeness of that information is not easy to assess. Compared to official registered migration data, it can be assumed that we obtained information about approximately a half of the emigrated population (*ca* 30,000 people).

Ethno-cultural characteristics

Ethno-cultural characteristics do not include any core characteristics. However, in Estonia, information was collected about the following characteristics, based on tradition and the stakeholders' requests: **ethnic nationality, mother tongue, command of foreign languages, ability to speak an Estonian dialect (local form of language)**, and **religion.**

Ethnic nationality has been included in all previous population censuses organised in Estonia, and it is a characteristic of great interest. The pilot census of PHC 2011 also included a question about the second ethnic nationality, but it was dropped from the actual census. Ethnic nationality was also determined by using a classification. In principle, it was possible to refuse answering the question regarding ethnic nationality, but this option was virtually never used, as ethnic nationality is not a sensitive topic for the permanent residents in Estonia. Registers (Population Register, PHC 2000) were used to supplement the data on ethnic nationality for 1% of the respondents; the question remained unanswered for 0.1% of the respondents.

Mother tongue is one of the most important characteristics that identify Estonians, and it has been asked about during previous censuses as well (with minor differences in wording). This characteristic is not a sensitive one in Estonia, but sparks great public interest. A classification was used here as well. Registers (Population Register, PHC 2000) were used to fill in the gaps in 1.7% of the cases, and the value of this characteristic is missing in the database in only 0.1% of the cases.

Although in a different wording, the question about the **command of foreign languages** has been included in several censuses, but it has mostly been limited to a prescribed list or restrictions (e.g. languages of the ethnic nations of the Soviet Union). In this census, the command of foreign languages was asked with regard to all persons aged three or older. Compared to earlier censuses, there were more options for indicating the languages, including a classification and a free text field. As requested by philologists, each respondent could indicate up to twenty foreign languages. As a result, a total of 127 foreign languages and 159 native languages were indicated. Curiously, there were instances, where a respondent had added programming languages under foreign languages, but they were deleted during the encoding process. Information about the command of foreign languages is not available for 1.9% of the respondents. No sources were used to supplement the results.

For the first time, the census included a question about the **ability to speak a dialect**. It was added to the questionnaire at the request of the ethnic regionalists of South-eastern Estonia (activists of the Seto and Võro movement)[a]. The question was asked from all persons aged three or older who spoke Estonian as their mother tongue. This question remained unanswered by 1.8% of the respondents; the information was not supplemented.

Religion is a question, the necessity of which has been subject to discussion before every census organised by the Estonian state, although it is not a sensitive topic for Estonians. Instead, the main issue has been the importance – and hence, the necessity – of this topic for the society. PHC 2011 was no exception to that rule and, similar to PHC 2000, the questionnaire included a question about the religious affiliation of the permanent residents of Estonia. Two religion-related questions were asked from respondents aged 15 or older: whether the respondent had any religious affiliation (answering was voluntary), and what that religious affiliation was. Similarly to other characteristics, respondents were given several options: firstly, multiple-choice answers, secondly, a classification, and thirdly, inserting free text (the name of the religious denomination) on the designated field. The first question remained unanswered by 2.2% and the second by 0.6% of the respondents.

Health and disabilities

One issue occasionally considered in censuses relates to disability, state of health and coping. However, such questions were not included as core topics of the 2010 census round. Formerly, Estonian censuses contained questions about having a certain disability (blind/deaf/mute/mental disability). During the previous census in 2000, people were asked whether they had an officially registered disability (invalidity). At the request of the Ministry of Social Affairs, the National Institute for Health Development, and demographers, the PHC 2011 questionnaire was supplemented with two health-

---

[a] Indrek Jääts, Lugege meid üle, Akadeemia 25, No. 6, pp. 1076–1110.

related questions: one about the presence of a **long-term illness or health problem** and the other about **limitations of everyday activities due to health problems**. The information gained from the second question was also used for estimating disability-free life expectancy. These questions were unanswered by 1.8% and 1.9% of the respondents, respectively, and no other sources were used to supplement the results.

Questions about household and family

"Household" and "family" are basic concepts of the census. Household represents an object directly subject to enumeration, family is defined through household. PHC 2011 used the following definitions.

Household consists of a group of people usually living in a common dwelling, who share available household facilities (common budget and food); a person living alone is also a household.

Households were divided into private households, institutional households and households of homeless persons.

- Private household – a household living in a dwelling. A private household does not include persons who at the moment of census lived permanently in an institutional household.
- Institutional household – a household consisting of persons who at the moment of census lived in an institution that operates 24 hours a day throughout the year, and provide maintenance for persons living there – accommodation, food and, if necessary, care and treatment (e.g. substitute home, care home).
- Homeless – a person who did not have a place of residence (dwelling) at the moment of census, i.e. the person slept in random cellars, staircases, boiler rooms, abandoned buildings or in shelters for the homeless that do not provide 24-hour accommodation.

Family (family nucleus) – family is defined in a narrower sense (family nucleus), consisting of persons who are living in the same household and are related as husband and wife, as cohabiting partners or as parent and child. A family nucleus can thus be:

- a legally married couple or a cohabiting couple without children (couples without children);
- a legally married couple or a cohabiting couple with children (couples with children, children need not be common);
- a lone parent with a child/children.

Each household member can be a member of only one family nucleus. A family nucleus cannot consist of more than two successive generations. If the household comprises three or more generations with a parental relationship between them (such as a child, his/her mother and grandmother), the family nucleus is formed of the two youngest generations.

The mandatory information about a household consists of the list of household members and information about **all relationships between the household members**. The best format for presenting all relations is a (triangular) matrix, containing the relationship between each couple of household members. The list of potential relationships is the following: spouse (only legally married), partner, child (only a biological or legally adopted child), spouse's or partner's child (not a biological child of the person), mother or father (including the adoptive parent), mother's or father's spouse or partner, sister or brother (including half-sister, half-brother), grandparent, grandparent's spouse or partner, grandchild, spouse's or partner's grandchild, other relative (including a great grandparent, great grandchild, father-in-law, mother-in-law; also the spouse's or partner's relative), a non-relative.

In Estonia, this type of a relationship matrix was used for the first time PHC 2011. The accurate presentation of a relatively complex relationship matrix was ensured by a large amount of (strict and soft) logic checks that excluded describing households with an illogical relationship pattern (e.g. the father is younger than the son, one person having several partners, etc.).

The relationships between household members were used to derive core characteristics about the family nucleus (**type of family nucleus, members of the family nucleus**), households (**type of private household** and **number of members in the private household**) and persons (**status in household and status in family**). The relationships between household members were supplemented to the extent of 1% based on the data of the Population Register.

The relations between household members were used to derive the following household characteristics:

- ratio between legal marital status and de facto partnership;
- households consisting of same-sex partners;
- households with a missing generation (households with grandparents and grandchildren, but no parents);
- reconstituted families (at least one child is a non-common child);
- multi-family households (households with more than one family nucleus);
- households with several generations (multi-family households consisting of at least three generations).

The relation between a household and a dwelling is described by the core characteristic **tenure status of dwelling**.

Agriculture

The questions concerning agriculture are not core questions. In Estonia, only one question was asked about agriculture, which ascertained whether a household produces any agricultural or horticultural products or keeps farm animals (including apiculture) for own consumption. Obviously, this did not include pets. Although a much more extensive questionnaire was initially considered, it was discarded after the pilot census. The question added to the census questionnaire gives an opportunity to conclude a further selective survey in the households that are engaged in horticulture, agriculture or cattle-farming.

Agriculture is not shown in the census tables, as this topic belongs to the Agricultural Census.

Dwelling and living conditions

There were 14 questions concerning the dwelling in the list of core characteristics. All these questions were indeed asked in Estonia, and very little was added to them. For example, the households without a bathroom were asked about the existence of a sauna as a washing facility, based on the traditional Estonian way of life. The core questions in the section dealing with dwellings contained the following: **type of household** (either a private household, institutional household or homeless), **type of dwelling**, **occupancy of dwelling**, **location of dwelling** (in the case of occupied households coincides with the occupants' permanent place of residence), **owner of dwelling**, **number of occupants** (a derived characteristic), **total floor area of dwelling and/or number of rooms**, **occupation density** (a derived characteristic), **piped water supply**, **toilet**, **washing facilities**, **heating system**, **type of building and time of construction**.

Although, in order to meet the requirement of core characteristics, it was necessary to ask about either the number of rooms or the total floor area, both were asked about in Estonia. Additionally, the households without a bath or a shower facility were asked about the existence of a **sauna** (it is not included in the list of internationally recommended characteristics). As for the recommended but voluntary characteristics, the PHC 2011 questionnaire (as in previous censuses) contained a question about the **existence of a kitchen or kitchenette**.

Additionally, in Estonia, a question was asked about **unoccupied or temporarily occupied dwellings owned by the household.** If such dwellings existed (could have been more than one), the respondent had to fill in a full dwelling questionnaire about them, including the exact address of the dwelling and its location on the map. Here it is important to keep in mind that this category did not involve summer cottages (their inclusion would have significantly increased the number of temporarily used dwellings), but only the dwellings that allowed normal use all year round (presuming the existence of a heating source).

In comparison to PHC 2000, PHC 2011 contained fewer questions about the amenities of the dwelling: questions regarding hot water, electricity, gas and sewerage were left out. The dwelling characteristics (including address) also provided a basis for deriving information about buildings with dwellings, including residential buildings, which are also included in the output tables.

The missing values of dwelling characteristics were substituted with information from two sources – the PHC 2000 database and the Register of Construction Works (EHR). The majority of dwelling characteristics received a good response, missing values constituted only 1–2%. More detailed information was most frequently missing in the case of unoccupied dwellings.

Table 2. Existence of values for dwelling characteristics and substitution of missing values
(percentages)

Characteristic	Substituted value	Not in database
Type of dwelling	0	0
Occupancy of dwelling	0	1.74
Owner of dwelling	0	0.84
Time of construction	7.65	1.24
Number of rooms	0.77	0.49
Area of dwelling	0.36	0.51
Kitchen	0.70	0.47
Water supply	0.13	0.45
Washing facility	0.05	0.62
Sauna	0.19	6.59
Toilet	0.13	0.60
Heating	0.06	0.70

PILOT CENSUS AND LESSONS LEARNED

Purpose of the pilot census

Where possible, a pilot census is organised prior to the population census, in order to assess the census methodology and organisation. Pilot censuses were held already before the first population census was carried out on the Estonian territory in 1881; it was first mentioned in 1864.

In most cases, the pilot census is as similar to the actual census as possible, but it still has to be viewed as a lesson and everything that went wrong during the pilot census needs to be improved. The pilot census of PHC 2011 was carried out exactly two years before the actual census. Exact timing ensured the presence of equivalent weather and other conditions. The two-year interval provided enough time to implement the conclusions reached after an analysis of the pilot census (PL) results.

Since the purpose of the pilot census is to check the census methods, instruments and the organisation of work, the aim of processing the data collected is not to estimate the values of measurable characteristics. As a result of data collection, questionnaires are checked for questions that are incomprehensible or difficult to answer, software malfunctions and other problems interfering with or endangering the census, which should be avoided during the real census.

There were several questions which the pilot census was supposed to answer, with the most important ones being as follows:

- Will self-enumeration work? Will people be honest or will there be a high level of tomfoolery and joking?
- Will the questionnaire work; are the questions and response options adequately understandable for the respondents?
- Is the questionnaire too long? How long does it take to complete it?
- How will people cope with online enumeration? How many people will opt for this method?
- How will the new data collection software, especially prepared for the census, function?
- How well do the register-based enumeration lists correspond to actual places of residence?
- What will be the main problems in connection with addresses, incl. writing them down?
- How successful is the establishment of spatial coordinates for dwellings?
- How successful is the monitoring of interviews with the special software?
- Which problems occur in the activities of enumerators and in the management thereof?
- Is the time reserved for the primary data processing stage sufficient for updating the enumeration lists to the required extent?
- How will the census management system work?
- How will the interview data be transferred to the database?
- What is the attitude of the people and the media towards the census?
- How high is the refusal rate and what are the reasons?
- How will the census support systems work?
- Is the schedule adequate?

During the pilot census, both the respondents and the enumerators were asked to provide various additional information (including in text format): the respondents and enumerators could add their comments to the questionnaires, and this option was actively used by online respondents; after the census, the questionnaire was filled in by enumerators, and the respondents were also involved in a feedback survey.

Timing and schedule of the pilot census

The greatest problem before starting the pilot census was to create a new software application, to describe the questionnaires in the definition module and to test the instruments, all of which took longer than anticipated. The software was not ready and sufficiently tested by the time when the pilot census was initially supposed to start (October 2009). In order to maintain the intended two-year interval between the test and the actual census, i.e. to conduct the actual census under the equivalent weather conditions as the pilot census, it was decided to postpone the critical moment of both the pilot census and the actual census, with the term finally being set on 31 December at 00:00 in 2009 and 2011, respectively. During the pilot census, the three-stage time schedule was also tested for the first time, the schedule consists of the following stages: the e-census, primary data processing and census interviews.

The census moment started a 53-day e-census period (31.12.2009–21.02.2010), followed by 11 days (22.02.2010–4.03.2010) of primary data processing. This involved removing from the enumeration lists all the dwellings, the occupants

of which had duly completed all the required questionnaires (dwelling, household and all personal questionnaires). This was followed by laptop-assisted interviews, which were carried out from 5.03.2010 to 31.03.2010.

There was another proposal considered when preparing the pilot census schedule. Namely, according to one hypothesis arising in the course of discussions, Estonia has a well-established tradition of filing income tax returns online, and therefore, it was believed that there could occur a certain synergy between the two online procedures – filing income tax returns could also push people to perform an online self-enumeration. There was even a thought of linking the filing of income tax returns with completing the census questionnaire, but it proved to be impractical for various reasons. The connection with filing income tax returns was assessed by feedback questionnaires and it proved to be extremely weak (the existence of such a connection was only confirmed by 2% of the respondents); hence, the idea was left out from the actual census.

Regions of the pilot census. Problems with addresses

The regions of the pilot census were determined so that they would be as diverse as possible – the ten regions for the pilot census (25 enumeration areas) included cities of various sizes, a town, a small town, and a low-density rural municipality, regions with expected rapid increase near a big city and cities with a decreasing population, Estonian-, Russian- and mixed-language areas in Northern, North-Eastern, Southern, Western and Central Estonia. Thus, the pilot census was intended to cover a significant share of problems related to population distribution, and to a great extent, it was indeed achieved. Certainly, there was no intention of having the population of the regions included in the pilot census to represent the entire Estonian population, and hence there was no point in comparing the single characteristics measured during the pilot census either with regular statistics, surveys or data from actual censuses.

The main problem that occurred in the pilot census was associated with addresses. Estonia had already taken a few steps towards organising address data – e.g. by developing the Address Data System (ADS), which should have facilitated address use, but unfortunately, it had not yet been introduced to all agencies and was also unfamiliar to ordinary people. The enumeration lists were prepared by using addresses from the Population Register, supplemented by addresses from the Land Board and, ultimately, from Estonian Post. In the end, more than 7,000 addresses were established as the places of residence of *ca* 10,000 persons (approximately one and half times more than expected) – this fact demonstrated that the addresses from different sources were not compatible, there were duplicates among them, and besides residential buildings, the address list of dwellings also included non-residential buildings (e.g. substations or warehouses). During the census, in the newly-built areas, enumerators also found such dwellings that were not listed anywhere.

Although intensive work with addresses continued also after the pilot census, it has to be admitted that addresses caused the most problems during the actual census as well. It was also to be expected in view of the following facts.

▪ This was the first time when the respondents had to indicate their address by themselves – previously, the address of the dwelling had been indicated by the enumerator and it usually coincided with the predetermined address on the enumeration list.

▪ The address was marked by using the new ADS-system (introduced in 2008, the term for implementation was in 2013), which had not yet been fully implemented and was somewhat different (in form and content) from the short addresses formerly used as postal addresses.

▪ People still think that registering their place of residence in the Population Register is voluntary (as it was in 1994–2000) and despite the fact that the registration of the actual place of residence is now obligatory, the registered place of residence differed from the actual place of residence in case of approximately one-fifth of the population. This was also reflected in census data, where, in spite of notification, some people still indicated their registered place of residence instead of their actual place of residence.

▪ The benefits provided by local governments (allowances, kindergarten places, etc.) cause people to register their place of residence in the cities or rural municipalities that offer better benefits, although their actual place of residence is elsewhere.

▪ Marking the dot indicating their dwelling on the (settlement) map displayed on a computer screen was beyond the capabilities of quite many people, because they had not performed such assignments before.

Lessons learned from the pilot census

The pilot census provided answers for the majority of the previously formed questions. Thus, it became apparent that self-enumeration worked (even better than expected); the quality of self-enumeration results was not worse than that of the census interview results; and tomfoolery was virtually non-existent. Both the media and the general public were

moderately positive about the census, but there was some criticism in social media. The pilot census did not gain any special attention from the media (and it was not the intention).

The most criticism and proposals were voiced about online questionnaires. The questionnaire was considered too long and its completion too time-consuming (in part, this concerned the software performance of the online census). The questions regarding data that were available in registers were considered redundant; i.e. it was recommended to use pre-filling. There were questions that people did not want to answer (health status, the exact address of the secondary place of residence and place of work, some information related to the place of work) or did not know how to answer (the grandparents' place of birth). It also became apparent that there was a surprisingly poor connection between the people actually living in the region and those entered into the enumeration lists (Figure 2). Among other things, this was caused by the selection of pilot census regions (as it later turned out, Peetri village in Rae rural municipality was the fastest-growing settlement in the entire Estonia in that period). In any way, it was a sign of danger – reaching the respondents based on the enumeration lists could pose a major problem.

Figure 2. Permanent residents entered in the enumeration lists of the pilot census and those actually living in the region
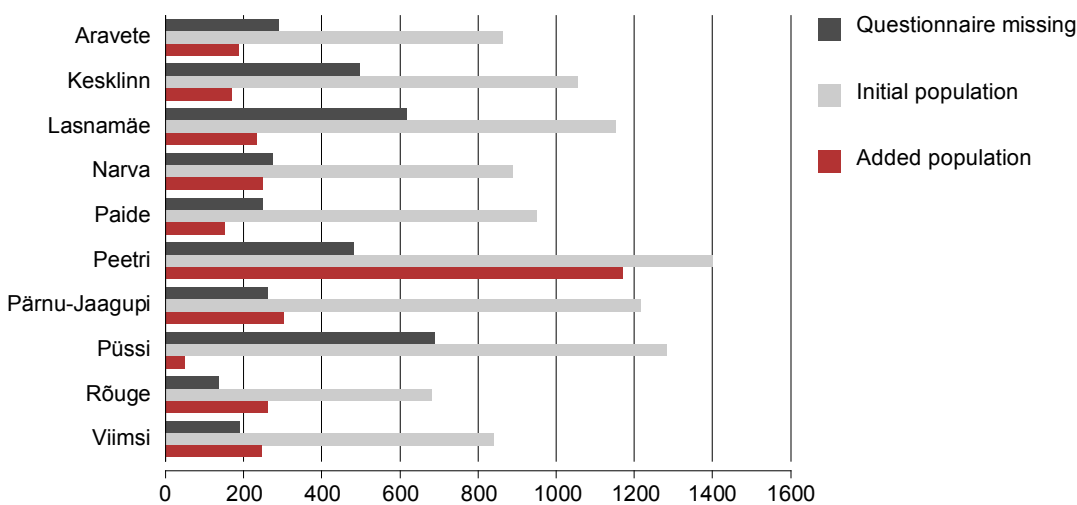
The enumerators' greatest failure was that in the case of some dwellings it was not ascertained whether they were occupied or not, such dwellings amounted to approximately 10% of all dwellings. In order to avoid this problem during the actual census, the organisers reworded the questionnaire, supplemented the training programme, and specified the enumerators' rules of conduct. These measures worked and the above-mentioned problem was virtually non-existent during the actual census.

There were also shortcomings in the software operation, but no major system collapse took place. Some of the problems were solved already in the course of the pilot census.

Time required to fill out census questionnaires

One important objective of the pilot census was to find out the time required to fill out the questionnaires, to be considered in estimating the resources intended for the census. For that purpose, the time required to fill out the questionnaires was automatically registered (except the time spent on entering the census environment). The time required for filling out census questionnaires during the pilot census is shown in Table 3.

Table 3. Average time required for filling out questionnaires during the pilot census
(minutes)

Questionnaire	Online	Interview
Dwelling	4.93	3.82
Person	16.09	7.54
Household	10.11	5.16
Annex to household questionnaire (in case of interview, included in household questionnaire)	4.52	2.6
Entire object (dwelling, household and 2.3 persons)	56.5	26.3

It appeared that the time required to fill out the set of questionnaires was less than an hour in the case of an average-size household (2.3 members) during the pilot census, but less than half an hour when being interviewed by an enumerator. The enumerator's work speed improved significantly in the course of the census. To some extent, the results depended on software specifics and corrections made in the course of the census.

Changes in methodology and organisation of work based on the results of the pilot census

The modification of the schedule was one of the most significant changes made based on the results of the pilot census. As the enumeration activity dropped significantly in the middle of the e-census, it seemed practical to shorten the time scheduled for the e-census to one month. The hypothesis that the partial overlapping of the e-census with filing the income tax returns would increase e-census activity was not proved (this was verified by interviews), thus there was no need to extend the e-census to February, when filing the income tax returns started. The data processing period between the two stages of the census was extended to two weeks and the interview period to a month and a half. The overall length of the census (3 months) remained the same.

In order to ensure a more consumer-friendly census environment, a series of proposals was submitted to the software developers, and minor inaccuracies that had to be removed were also recorded. The greatest substantial problem related to software was the deletion of one data set (descriptions of household relationships) associated with approximately 40% of the respondents of the interview census, which was luckily the only data loss that occurred during the pilot census. This error was fixed.

Predicting the participation rate of the e-census to assess the need for census resources. A model-based estimation was made regarding the e-census participation rate, but it turned out to be overly cautious. This prediction was necessary to assess the need for enumeration areas/enumerators, but also to estimate the necessary capacity of communication channels during the e-census. As the overestimation of the importance of e-enumerators posed a greater problem, which would have led to hiring and training too few enumerators, a rather conservative estimate was made – it was assumed that the participation rate in the e-census would be *ca* 25%. It was decided to determine 2,000 enumeration areas with 2,000 enumerators. The enumerators with supervisors and managers constituted a hierarchical system (3 regions, 15 districts, 123 supervision areas, 2,000 enumerators).

Changes in the census questionnaire

The census questionnaire was subject to quite a number of changes, aiming at shortening the questionnaire and facilitating its completion. Some questions were removed – e.g. those regarding the second ethnic nationality, the second citizenship, age at the birth of the first child, the average number of working hours per week in total and in the main place of work, the reason for staying in the second place of residence, and the second source of subsistence. In the case of some questions, the age limit for the respondents was increased – questions about education and the source of subsistence were asked from respondents aged 15 and older (during the pilot census the age limit for questions regarding education was 10, and the source of subsistence was asked about from all respondents).

A rather thorough reformulation was conducted with regard to the questions on education. The accuracy of answers to questions concerning education was verified by comparing the pilot census data with the level of education recorded in EHIS (the Estonian Education Information System), available for people who have graduated from Estonian educational institutions within the last five years. It occurred that quite many people had been incorrect in identifying their level of vocational education. In order to avoid that, the wording and order of the questions was altered and in the case of one set of questions, the respondents' age was also taken into account (avoiding an unreal response being indicated).

There were further simplifications made in the personal questionnaire. During the actual census, the address of the secondary place of residence and the secondary place of work was to be indicated at the settlement level (during the pilot census, it was determined at the dwelling and house level), and in terms of questions on the place of birth and migration, the indication of *krais*[a] and *oblasts*[a] of Russia was discarded, and the month of changing the place of residence as well (instead, only the year was asked).

From the household questionnaire, the set of questions on small-scale agriculture (additional household questions) was removed and replaced by a single question, which allowed ascertaining whether a household grows any food products for own consumption or not. Based on this information, a further selective survey can be conducted later on.

The dwelling questionnaire was also simplified: the questions on hot water supply and type of sewerage, as well as a few additional questions were dropped. The existence of a sauna was asked only from households without a bath or a

[a] type of an administrative unit

shower facility. The option of giving two answers for one question (type of heating, owner of dwelling) was also discarded, as it created further technical problems.

Initially, it was hoped that several questions could be pre-filled based on registers, which would have shortened the time spent on completing the questionnaire. Due to various technical reasons, however, pre-filling was not available for testing during the pilot census, but the idea of pre-filling was still not dropped. Some questions gave grounds to rather heated discussions after the pilot census. Such questions concerned mostly health restrictions and the grandparents' country of birth (the latter was not tested during the pilot census, but it had been used in a survey). Due to pressure exerted by the stakeholders, these questions were kept in the census questionnaire.

Some questions were also added to the census questionnaire after the pilot census. One such additional question concerned the ability to speak a dialect, which was asked about from all respondents, aged 3 or older, whose mother tongue was Estonian. Unlike the pilot census, the unemployed were asked about the information on their last place of work (occupation, economic activity of the place of work, employment status, and the last time of employment). Another set of questions was added to the household questionnaire. After enumerating the household members and the temporary residents of the household, a question was asked about close relatives who had left Estonia within the last twelve years and stayed abroad. Besides identification, information was also requested about the year of leaving Estonia and the country of permanent residence of such persons.

Assessment of the success and quality of the pilot census

During the pilot census, several problems and bottlenecks related to software and organisation were discovered and mostly removed. Certain changes were made to the questionnaire and the census schedule. In that sense, the pilot census achieved its goal.

Assessing the quality of the pilot census required an evaluation of its **coverage**. For the pilot census, the organisers had developed and the PHC board had approved quality criteria, which prescribed coverage and the rate of e-census participation corresponding to very good, good and satisfactory quality standards.

Coverage is one of the crucial criteria of the census quality. Coverage shows the ratio of actually enumerated items to items that were subject to enumeration.

- N – number of items subject to enumeration (general population)
- L – census population, i.e. the number of enumerated items
- K – **coverage rate** (K = L/N)

Coverage rate is often expressed in percentage. K < 100% stands for **under-coverage**, i.e. some of the items subject to enumeration have not been enumerated. K > 100% stands for **over-coverage**, i.e. enumeration includes items that do not belong to the general population or a certain amount of items have been enumerated more than once. The extent of under-coverage and over-coverage is also measured in percentage:

- under-coverage rate is (N–L)/N,
- over-coverage rate is (L–N)/N.

Table 4. Quality criteria for the pilot census
(percentages)

	Very good	Good	Satisfactory	Poor
Coverage rate				
Dwellings	99	97	94	< 94
Persons	95	90	80	< 80
Online enumeration rate				
Persons	≥ 18	≥ 13	≥ 8	< 8

The coverage of the pilot census was very difficult to evaluate. The greatest problem was the determination of the general population N, especially in the case of dwellings, because occasionally there were rather remarkable discrepancies between the initial sources and the data collected by the enumerators. All in all, the coverage of dwellings was deemed "satisfactory", the coverage of persons "good" and participation in the e-census "very good". The coverage of single characteristics was in most cases (*ca* 99%) very good, and the answers were almost exclusively reliable. However, a number of problems were revealed by the comments – more than five thousand were received.

The persons residing in the pilot census areas were notified of the opportunity to participate in the census online, but besides that, all persons with an Estonian personal identification code could fill in the census questionnaire online, using either an ID-card or the access codes of major Estonian commercial banks (SEB, Swedbank) to enter the census

environment. Such voluntary respondents, amounting to over 5,000, provided significant value added to the census team, primarily due to the comments added to the questionnaires.

The second pilot census – a mini pilot census

As both the questionnaire and software underwent significant modifications after the pilot census, the instrument needed another test. For that purpose, a simulated data set was generated according to the test plan, including different types of objects and persons, and prescribing various problematic data combinations and situations. This 'mini pilot census' took place in September 2011. The participants of the mini census included the staff of Statistics Estonia and volunteers, who played the roles prescribed by the pilot plan.

The analysis of the results of the mini pilot census indicated that a significant share of former insufficiencies had been rectified, and there were no deletions or data loss. Thus, the basic preparatory works were complete.

Establishment of enumeration areas

One of the most labour-intensive preliminary works for all censuses is the establishment of enumeration areas, i.e. distributing the entire enumeration territory between enumerators so that no dwelling would be included in the list of multiple enumerators or get left out. Even in Estonia, with its limited territory and small population, these are huge tasks – for instance, in 1922 and 1934, a total of 17,000 and 18,000 volunteer enumerators were used, respectively, meaning that each enumerator had to write down the information regarding an average of 60–65 persons and a couple of dozen households and dwellings. They had one or two days to do it, but work started already a few days before the census, and the results were verified once more on census day.

Virtually before each census, complaints have been made about the poor quality of the maps and plans used for the census. Prior to the census of 1922, it was noted that the situation was particularly bad in the cities, where entire districts were missing from the maps. In Soviet times, there was another problem – the maps were there, but they were intentionally distorted. Today, the quality of maps is remarkably better than earlier, but some problems persist – due to rapid progress, the maps based on aerial photographs made for the previous census do not meet the requirements set by the modern GPS-system. This means that the spatial coordinates of residential buildings are indicated on new maps and in new ways, and thus specification had to be made during the census – e.g. adding new buildings to the maps.

In principle, there are two options for determining enumeration areas: either based on settlement borders and the street network (as done previously) or, on the map, designate areas with an optimal shape, taking into account the natural environment (e.g. rivers), not only settlement borders. The second option was chosen for the 2011 census.

There were as many enumeration areas as there were enumerators, 2,000 enumerators were planned to be used for the 2011 census (the actual number was reduced by a few dozen). This meant that an average enumeration area included 600–700 persons and about 300 dwellings (some of them unoccupied or used on a seasonal basis). The workload is different in the city and in the country: in the city, an entire enumeration area could fit in one apartment building, whereas in the country, it could comprise a dozen or more villages. Therefore, population density was also taken into account when establishing the enumeration areas. The number of persons subject to enumeration was lower in low-density areas and higher in high-density areas. Based on the addresses of dwellings belonging to an enumeration area, the number and list of persons registered in each enumeration area and each dwelling was ascertained by using the data from the Population Register. Such preliminary work resulted in rather precise enumeration lists for the enumerators, containing more or less all dwellings (addresses and spatial coordinates) belonging to the enumeration area and the residents registered there.

E-CENSUS

Access to the census environment and security of the e-census

As planned, the Population and Housing Census 2011 started with an online e-census in the morning of the last day of 2011. The census environment could be entered using an ID card, bank access codes or Mobile-ID. Using the bank access codes proved to be the most popular method, as online banking services (incl. filing income tax returns) have been in use in Estonia for years already.

The identification of persons logging in to the census environment was protected with thorough security measures. The only instance that could be regarded as a security leak, which was also reported by the media, occurred on the third day of the census (the first business day after the census moment) when there were a couple of situations where the identification data of the enumerator (not the census questionnaire itself) remained visible to a third party. This potential security leak was taken very seriously. A high-level committee was assembled for investigation, which revealed that the fault did not result in any actual data leaks. After the discovered flaw was eliminated, there was no possibility of unauthorised access to any census data – the strictest security requirements were met in this census.

Support systems of the census

The use of online self-enumeration required efficient support for any technical issues associated with login (authentication), the questionnaire or any other areas. During the census, the participants had the opportunity to contact the support team by telephone or e-mail with any questions they had about the methodology or any technical issues with the computer. The answers were generally quick; the support team was overloaded only during the peak period (2 and 3 January 2012). All help requests were registered. In total, the support team consisted of 16 consultants working in two shifts; they were also aided by the employees of statistics departments. During the e-census, the support team was contacted 50,503 times by telephone and 7,668 times by e-mail. The largest number of questions concerned technical problems; in terms of methodology, the concept of 'household' proved to be the most difficult one to understand, particularly for Russian-speaking participants.

## Dynamics of participation activity during the e-census

To ensure the openness and transparency of the census, a separate website of the census was created and, during the e-census, it showed the number of persons enumerated and their percentage (%) of the official population in the country as a whole and in individual counties. The address of the website was widely publicised in advertising materials and on posters. In order to manage the load on the census system, the site included a three-colour counter measuring the number of respondents who were currently logged in to the census environment – the pointer was in the green sector if the number of online respondents was well below the rated capacity of available communication channels; it was in the red if the number of respondents approached or exceeded a critical limit; a situation between these two possibilities was indicated with the yellow sector.

The first and second day of the census were calm and the pointer remained in the green sector. The most critical day was the third day of the census, 2 January, when the number of respondents exceeded expectations – despite the pointer constantly being in the red. The time required for responding was doubled compared to the previous days, exacerbating the congestion even further. The support team was unable to answer all the incoming questions and there was even a service interruption, which fortunately only lasted for half an hour. The capacity of data communications channels was increased after the failure and there were no further problems of this kind. The extremely intense start of the census, when *ca* 50,000 persons started to fill out questionnaires in one day, was followed by a rather quiet period of two weeks, with about 20,000 questionnaires being filled out per day (Figure 3, p 30).

The participation activity increased in the last ten days – partially due to additional notifications by e-mail as well as the competitive spirit that took over the respondents. Although the number of enumerated persons rose to 60,000 per day on the last days, there were no further setbacks. The capacity of the system had been increased enough to ensure that the pointer remained in the yellow sector even during the peak hours. As many questionnaires were still being filled out during the last days of January, it was decided to extend the e-census period by one more day, to include 1 February, and the census environment was closed on 2 February 2012 at 2:00.

As expected, the participation activity in the e-census was higher at the start and at end of the period and clearly lower during the time in-between. For instance, the number of persons filling out personal questionnaires was almost four times lower on 13 and 20 January than on 30 January. The day of the week was not particularly significant in determining response activity.

Figure 3. Personal questionnaires started and completed in the e-census by individual days
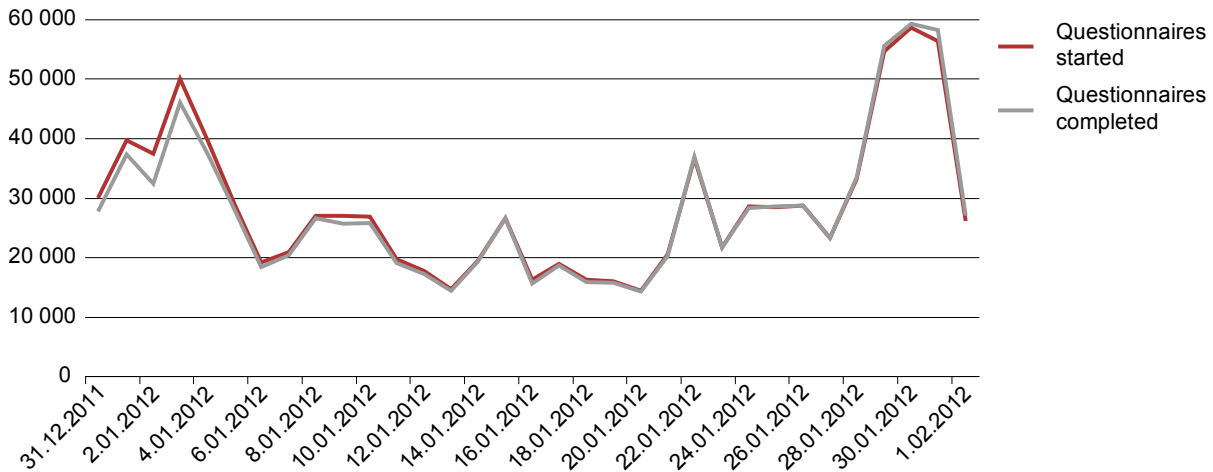
It was initially assumed that people would mostly fill out the questionnaires at work, i.e. during working hours. However, it turned out that this assumption was inaccurate – the peak period of filling out the questionnaires started at 18 o'clock and lasted almost until 22 o'clock. The people who filled out their questionnaires at work mostly used the afternoon hours between 12 and 16 o'clock. As expected, the number of people who filled out the questionnaire late in the night – between 3 and 6 o'clock – was extremely low (Figure 4).

Figure 4. Usage activity of the e-census environment over 24 hours
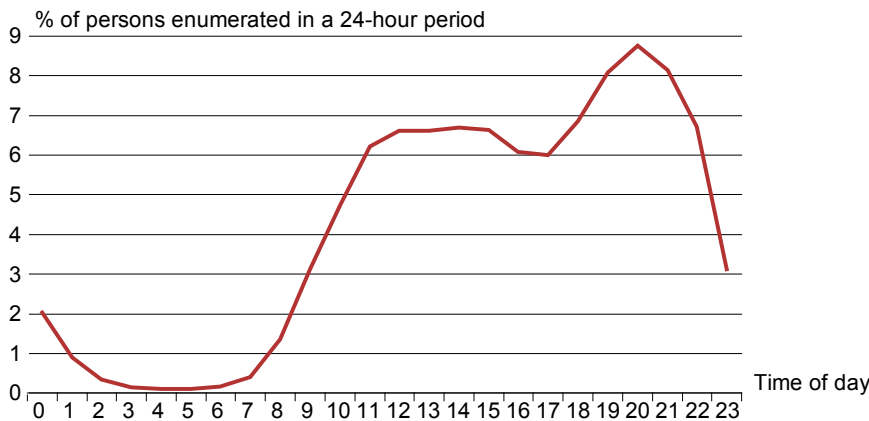
Persons enumerated in the e-census: number and distribution by counties

On 2 February 2012, the PHC website announced that a total of 815,933 persons, or 61.8% of the permanent residents of Estonia (according to the official population figures of Statistics Estonia, including migration), had been enumerated. However, this was not the whole truth. On the one hand, this figure did indeed show the number of personal questionnaires received, but it also included duplicates, which were estimated to amount to about 3% of the questionnaires. On the other hand, the count only included persons who had filled out all household questionnaires. After adding the persons who had filled out their personal questionnaires separately, without the household, the number of online personal questionnaires of permanent residents increased to 880,455, which constitutes 68% of the persons enumerated. A little over 1% of the e-census questionnaires were subsequently supplemented during the census interviews. In any case, approximately 67% of the permanent residents had participated in the census online. This exceeded all expectations. There were some expected differences between the counties but they were not large, which was another positive surprise. A comparison of Ida-Viru county and Tartu county, the counties with the lowest and highest participation rate in the e-census, respectively, only reveals a difference factor of less than 1.6 (Map 1).

Map 1. Participation rate in the e-census by local government units

Enumerated in e-census, %

9 – 49 (25)	
50 – 59 (84)	
60 – 69 (81)	
70 – 85 (36)	

— Border of county
— Border of rural municipality
○ City with municipal status
□ Rural municipality with an area smaller than 10 km²

50 km
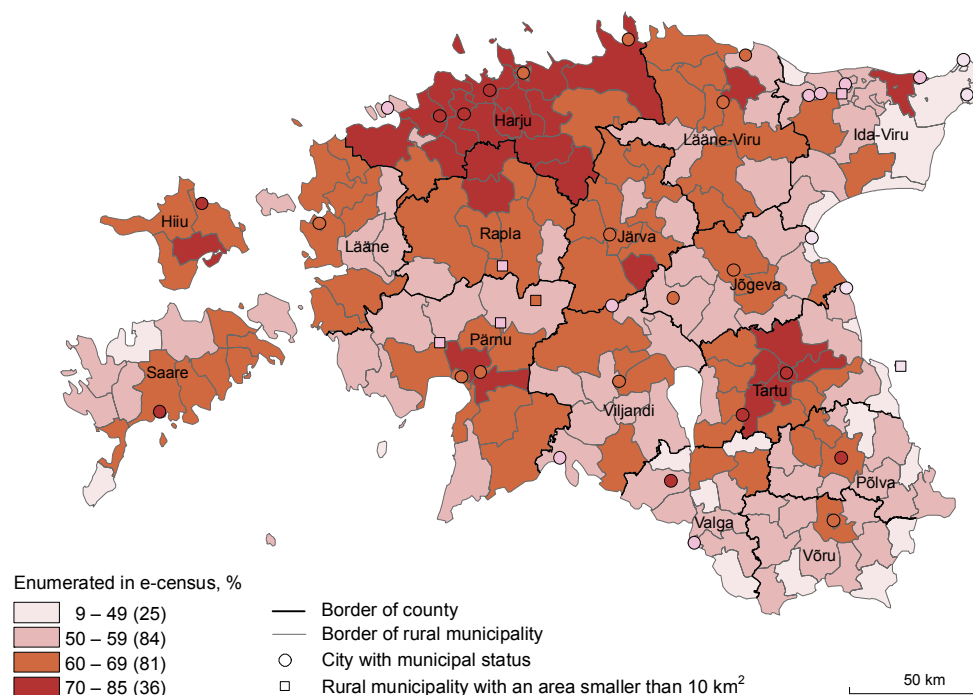
Who participated in the online census?

There were three main options for a person to be enumerated online: he or she could log in to the census environment (using an ID card, bank access codes or a mobile phone) and personally fill out the required questionnaires (usually the household, dwelling and personal questionnaires). The second option was for one person to log in to the census environment and fill out the household and dwelling questionnaires, after which other household members could fill out their own personal questionnaires without having to re-enter the census environment. As a third option, one household member could also fill out the personal questionnaires of all the other household members. As the person filling out a personal questionnaire had to be at least 15 years of age, this last option was used for children whose questionnaires were filled out by their parents.

While the e-census was generally used at a higher rate in younger age groups, the ratio of men and women among the persons enumerated online was rather similar to the ratio of men and women in the total population (Figure 5). However, a somewhat higher participation rate of women is a notable trend, since the subsequent analysis indicated that while the number of men slightly exceeds that of women in the younger age groups of permanent residents, women had a higher participation rate in the e-census.

Figure 5. Share of questionnaires received in the e-census by sex and age group
(duplicates have not been removed)
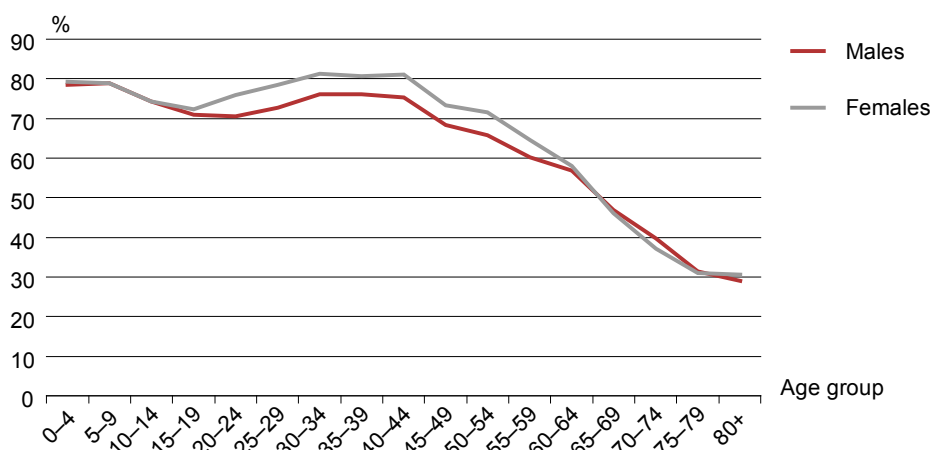
Figure 6 shows that most of the people whose personal questionnaire had been filled out online had done it in person. Of all the permanent residents aged 15 to 70 who were enumerated online, over 80% filled out their personal questionnaires themselves, and 55% personally logged in to the census environment. The youngest age group contained a relatively low number of those who logged in themselves, despite the expected high level of technical computer skills. It is likely that limited motivation was the decisive factor here. Women's higher participation activity – and maybe also better computer skills – is evident in the fact that women constituted 61% of the persons who personally logged in to the census environment. Among men who had their questionnaire filled out online only 38% had logged in to the census environment in person.

Figure 6. Share of persons who filled out their own questionnaire and logged in to the census environment in person among all personal questionnaires filled out online, by age group
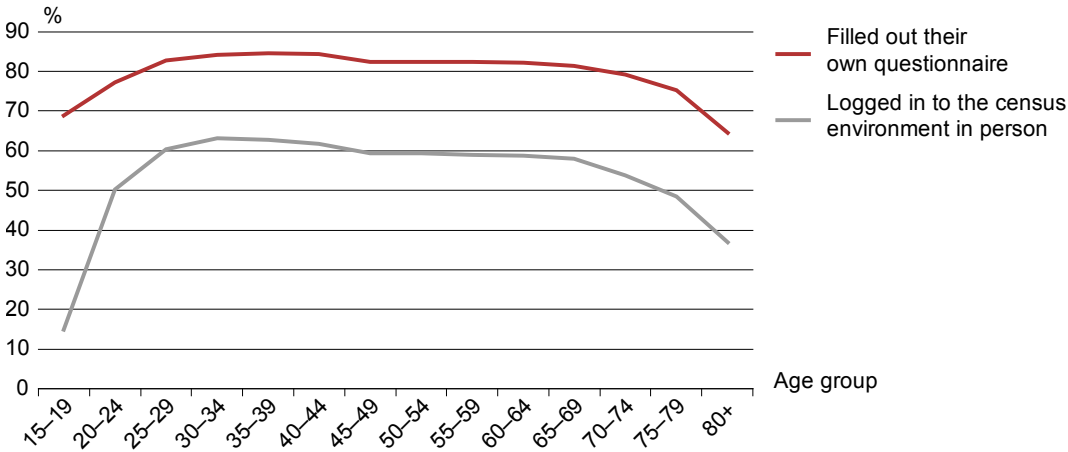
It can be concluded that men and women of up to 45 years of age feel confident working online – the share of those who personally filled out their online questionnaire starts to decrease after that age. The census was also supported by the network of public Internet access points, particularly the libraries, where online questionnaires could be filled out by those people who did not have a computer or an Internet connection at home.

CENSUS INTERVIEWS

Preparations for interviews. Initial organisation of data

While the enumeration lists were relatively accurate before the e-census, it was no longer the case afterwards, because nearly two-thirds of the people submitted their data online. If all members of a household had been enumerated, the dwelling of that household did not have to be visited (this point had been emphasised in the earlier promotion of the census) and the addresses of those dwellings had to be removed from the enumeration lists.

Data organisers started working as soon as the e-census was opened but they still ran short of time, because the number of questionnaires received during the e-census was more than double the anticipated amount. Furthermore, the duration of the e-census was extended by one day and this also reduced the time available for data organisation. This meant that the start of the census interviews had to be postponed for a few days in order to provide the enumerators with better-ordered enumeration lists. It was assumed that this would not create any major problems for the enumerators, as the number of persons to be enumerated after the e-census was around 200–250 per enumerator. However, the amount of work was not reduced by the same extent in the case of dwellings – the number of empty dwellings enumerated during the e-census was relatively low and the remaining dwellings had to be visited by the enumerators. They also had to verify the locations of the dwellings entered online. Nevertheless, the success of the e-census reduced the workload of the enumerators by at least a third, compared to the original estimates.

The work of the enumerators started with a training – the census interviews were preceded by a five-day course, presenting the overall ideology of the census, the questionnaire, the software and hardware, as well as the standards of behaviour and confidentiality.

Census interviews

The census interviews started on 20 February 2012 and lasted, as scheduled, until 31 March 2012, i.e. a total of 41 days. The interview period also included a public holiday – 24 February 2012. The enumerators were advised not to visit people's homes on that day, unless the person to be enumerated had requested it and a respective arrangement was in place. The census interviews were managed according to a hierarchical system – the country was divided into three regions, which were further divided into 15 districts and 123 supervision areas. The enumerators were free to choose their own working hours, but they had to adhere to certain limiting criteria to ensure the completion of work in an area within schedule. Adherence to these criteria was monitored by the supervision area and enumeration area supervisors in a monitoring system and the corresponding information was also sent to the PHC command centre in Statistics Estonia.

The task of each enumerator was to visit all the dwellings according to the enumeration list of the enumeration area. The list included the dwellings that had not been fully enumerated online (i.e., if enumeration was not completely correct regarding the dwelling or the household(s) living there or any household members). The enumeration of a person or an object was considered to be correct if the respective questionnaire had been filled out and marked as completed. Consequently, the enumerators had to collect the data on those dwellings, households and persons that were subject to enumeration but were not enumerated online or had questionnaires which were not marked as completed. The unfinished questionnaires had to be completed; if a question had already been answered, it generally did not have to be repeated (unless clarification was needed). In addition, the enumerators had to find any residential buildings or dwellings, which had been left out of the enumeration list, and then enumerate and enter them on the map. Furthermore, the coordinates of the enumerated dwellings/residential buildings had to be verified and corrected if necessary.

The general work procedures established that enumerators had the right to visit homes on a daily basis from 8:00 to 21:00 (incl. on weekends). A special arrangement was required for visits at any other time. This rule was generally observed. The greatest problem was establishing contact with some persons who were reluctant to cooperate. In those cases, the enumerators left letters or messages in the mailbox or at the door, made telephone calls, or asked neighbours for assistance.

Despite the major preparations made by the census team in compiling the lists of dwellings and persons by individual enumeration areas, it sometimes happened that the actual addresses differed from the addresses on the enumeration lists and the actual persons living at that address were not those who had been registered there.

At the end of each working day, the enumerators had to submit the data collected during the day to Statistics Estonia. This was the so-called synchronisation stage and it was the source of the highest number of complaints from the enumerators. Even though Internet access is available throughout Estonia, the capacity of the network (data transmission rate) was sometimes insufficient for transferring the PHC data sets (the data collected during the day) and the enumerators spent too much time doing this. On some occasions, the data arrived after a delay of several days, which was not a major problem in the end but created some confusion in the disclosure of information on the progress of the census.

Figure 7. Number of persons enumerated in the first 30 days (Stage 1) of the interviews, by day
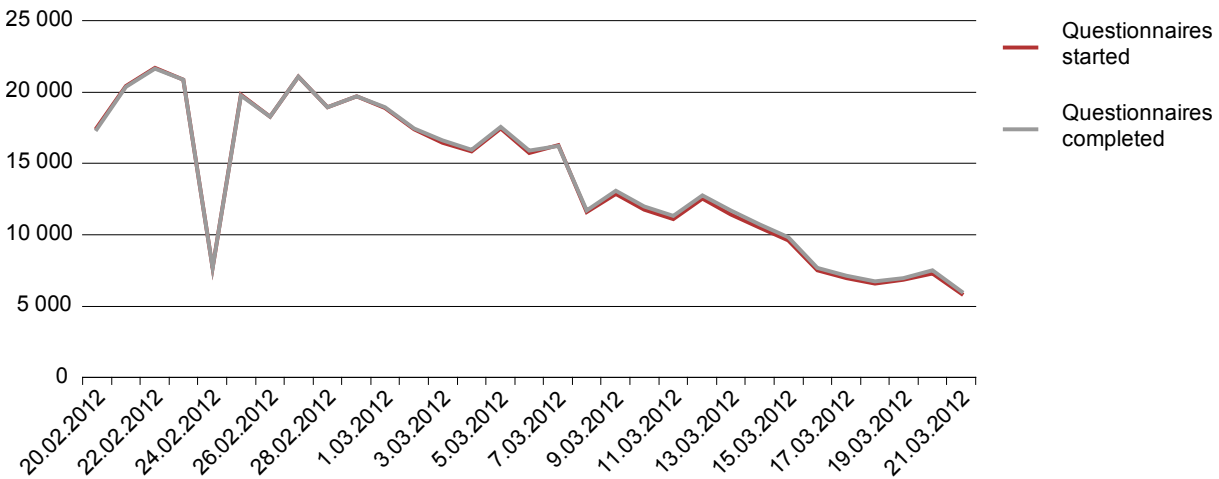
Updated results were also published during the period of census interviews, but there was an unfortunate technical fault, which meant that, for about a week, the figures published on the website were lower than the actual count. By 5 March, the difference between the actual enumerated population and the number of enumerated persons shown on the website was almost one hundred thousand, creating some concern among observant commentators. After the fault in the program was discovered and rectified, the published figure matched the information in the database but, due to connection problems, the incoming data flow was uneven, with a delay between the actual enumeration date and the receipt of data.

Figure 8. Persons enumerated during the interviews by the time of enumeration
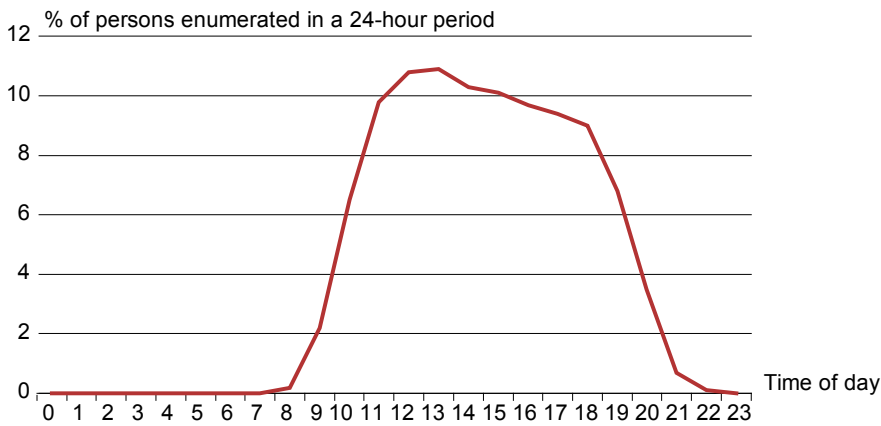
The number of questionnaires completed per day decreased gradually over the period of census interviews as more and more people were enumerated. Figure 7 also indicates that on the public holiday (24 February), the intensity of work was significantly below the usual level. Figure 8 reveals that the enumerators did indeed work during the prescribed hours. The fact that the middle of the day (from 11 to 15 o'clock) was the busiest period indicates that a large share of the persons visited during the census interviews were at home during the day and many of them were pensioners.

Second stage of the census interviews

After the majority of persons not covered by the e-census had been enumerated, the second stage of the census interviews was launched to clarify any confusing situations from the e-census and to find any persons previously missed by the enumerators. If the number of filled-out personal questionnaires of a household was lower than the number of persons on the respective list of household members, the enumerator had to go and interview the person(s) whose personal questionnaires were missing or unfinished. If a household questionnaire indicated that the dwelling included several households but there was only one household questionnaire for that dwelling, the enumerator, again, had to go and check whether the number of households was entered incorrectly (this was a surprisingly frequent mistake, with the

number of household members often entered as the number of households) or whether a household had been left out. A check was also needed for all dwellings which were vacant according to the owner. This was to clarify whether those dwellings included any residents (living there on a temporary basis or even unbeknown to the owner) who could otherwise have been omitted from enumeration.

This was the most difficult stage of the census interviews, because it created confusion among the respondents. Naturally, people did not remember that they had made a minor mistake when filling out their questionnaires. It was often also difficult to ascertain whether somebody lived behind locked doors and windows without curtains or not. Nevertheless, this census period resulted in the completion of almost 15,000 additional personal questionnaires, i.e. over 1% of all filled-out personal questionnaires. Even though many of the questionnaires filled out in this period were duplicates of previous questionnaires, they provided a large amount of additional information, significantly raising the quality of the census. As is the case with all censuses, there were some peculiar communication situations between the enumerators and respondents but, unlike in the previous census, the overall attitude towards the census was positive and no significant complaints were raised against the enumerators.

Were all the people subject to enumeration enumerated?

As the end of the census period drew closer, it became clear that each and every person who was subject to enumeration cannot be found and enumerated. This is a problem shared by all developed countries – people have become very mobile and are no longer tied to a single place of residence. It is quite possible for an enumerator to miss them, as they commute between several dwellings. We also no longer have (village) communities where everybody knows everybody and is able to provide information on their whereabouts. Quite the opposite, it is often the case that people who have lived in the same building and used the same main entrance for several years do not know each other, not to mention being able or willing to provide information about others.

There has also been an important change in the mindset – people are less and less inclined to disclose the particulars of their lives to others, especially the authorities (the enumerators were often seen as the representatives of the authorities). Such intransigent persons, refusing to be enumerated, have always existed – despite participation in the census being mandatory in Estonia – and their numbers have increased over time. Even though, in principle, the persons evading the census could be punished for breaching the law, this was generally not done, as the majority of them could not be found.

Map 2. Persons who were not enumerated

Persons not enumerated who contacted Statistics Estonia on 1–17 April, number by settlement units

228 (1)   55 (1)   10 – 28 (4)   • 1 – 9 (158)   HIIU   County

50 km
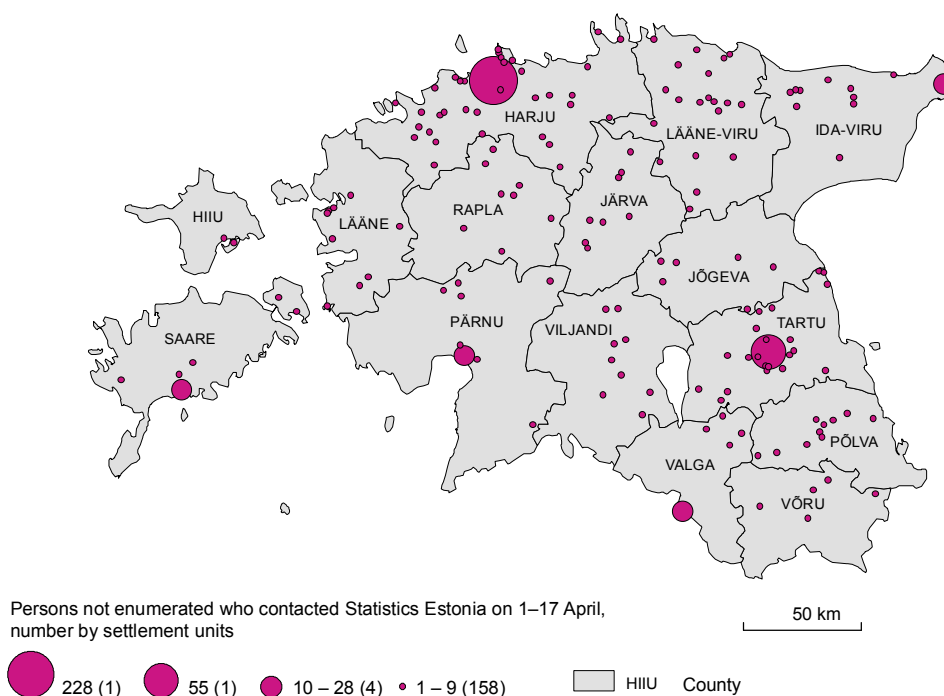
There were also some who were omitted from enumeration by accident. Below are a couple of examples. An old man called and said, "The Estonian state is not interested in me, I have not been enumerated." However, it turned out that the man's flat was enumerated and, according to the collected data, was inhabited by his daughter. The old man explained, "Yes, I registered the flat in my daughter's name. My daughter filled out the questionnaire on the Internet." The daughter had indeed entered her data online and indicated the father's flat as her dwelling (even though she should have entered her actual place of residence), but forgot to mention in the dwelling questionnaire that her father (also) lived in the flat. As a result, it was impossible for the enumerators to know that all inhabitants of that flat had not been enumerated.

Confusion between the actual and registered place of residence was probably one of the main reasons why some people were not enumerated. If a person had correctly filled out the questionnaire on the dwelling that was their registered place of residence, while their actual place of residence was located elsewhere, the enumerators did not visit that dwelling and any people who might have lived there were left non-enumerated (unless they themselves filled out their questionnaires online). There were also instances where a person refused to let the enumerator in. A journalist had a conversation with some men at a suburban shop in the morning of the first day after the census. "No, I have not been enumerated," a man declared. "They did knock on my door but I will not open my door to some random person. There were some papers in the mailbox but I never check my mailbox – who would want to write me?" Many people called in the final days of the census, saying that they had not been enumerated. Some of them were interviewed over the telephone, resulting in several hundred additional personal questionnaires. Nevertheless, it was clear after the census that a certain number of persons had been left out. This meant that the census had quality issues and extra work was needed to estimate coverage.

Map 2 indicates that omissions from enumeration were rather random, because the geographic distribution of persons who were left out correlates with population density: there was no region where the number of non-enumerated persons was disproportionately high compared to the enumerated population.

Internet or interview? The preferred method of enumeration depends on age

It was expected that younger respondents would prefer the Internet and older people would rather wait for a visit from the enumerator. However, it was interesting to see the turning point between the e-census and interviews in terms of age. It turned out that the age boundary was at 65 years – while 52% of 64-year-olds filled out their questionnaires online, the percentage dropped slightly below 50% among 65-year-olds.

Figure 9. Preferred method of filling out the questionnaires, by age
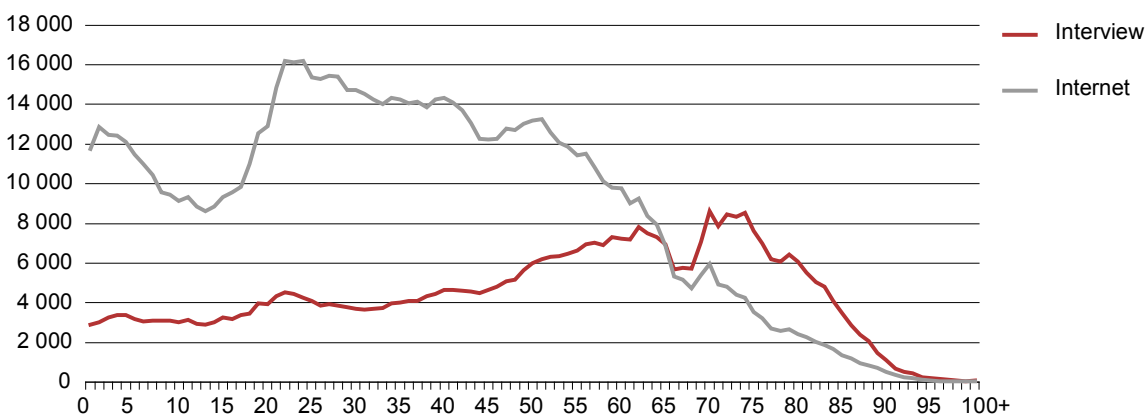
The relative importance of the online census was very high (77.3%) among children whose census questionnaires were filled out by their parents. The use of the e-census was particularly high in the case of younger children, starting from 80% for children under two years of age and dropping to 75% for teenagers. This result does not support the hypothesis that having school-age children increases computer and Internet use in a family. The percentage of online respondents was also high (79–80%) in the age group of respondents aged 24–33, and started to fall gradually after that. The percentage of online respondents drops below 75% for those aged 43 or over, then below 70% for those aged 50 or over, and below 60% for those approaching their 60th birthday. Among older people, the percentage of online respondents was around 30%.

Time spent on filling out census questionnaires

During the pilot census, the respondents (especially the volunteers) complained about the excessive length of the questionnaire, which made filling it out very time-consuming. Skilled computer users were also dismayed by the slowness of the application (software). After the pilot census, the census team worked hard to speed up the process of filling out the questionnaires by reducing the number of questions (even though some extra questions had to be added) and making the questionnaires more user-friendly through modifications in wording, the order of questions, and design. The software solutions were improved as well. The target was to reduce the time of filling out the questionnaires by 15%, but this target was achieved (and even slightly exceeded) only in the census interviews and not quite in the e-census.

Figure 10. Time spent on filling out personal questionnaires, by method and respondent
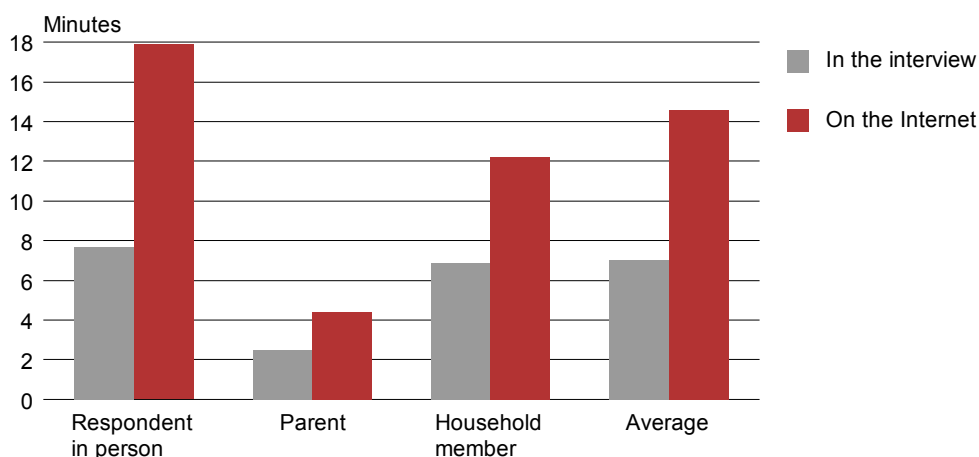
Figure 11. Time spent on filling out household and dwelling questionnaires
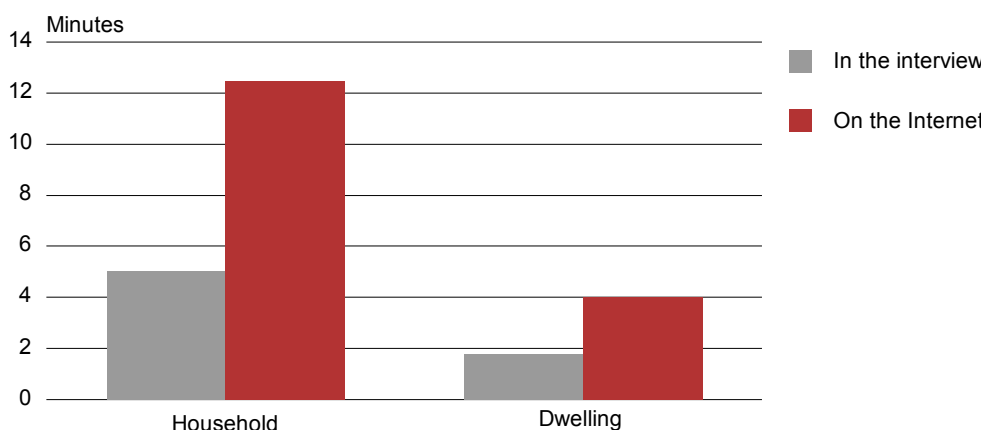
In both the pilot census and the actual census, the enumerators were able to work much faster than the persons who filled out their own questionnaires. The online respondents spent the largest amount of time on filling out their own personal questionnaires – this questionnaire was the first to be filled out and people had to read all the questions and many help texts. The time spent on filling out the questionnaire of a second household member was then reduced by a third, while the time parents spent, on average, on filling out their child's questionnaire was reduced by four times compared to the time spent on their own questionnaire. The final average time spent on filling out the online personal questionnaire was slightly under a quarter of an hour (14.6 minutes), which is approximately 10% less than in the pilot census. The overall trends were similar in the census interviews, but the numeric ratios were different. The time difference between entering the data of the first and second household member was only 10%, while filling out a child's questionnaire took three times less time than filling out the respondent's own questionnaire. Naturally, it was not always the case that one person provided the information on all household members – the adult members of the household (at least 15 years of age) usually gave their own responses or filled out their online questionnaires by themselves. During the interviews, an average personal questionnaire of a permanent resident could be filled out in seven minutes, which meant only 7% of time saved compared to the pilot census.

The situation was more or less similar with the dwelling and household questionnaires. The largest change occurred in the household questionnaire with the omission of additional household data, including a number of questions on small-scale agricultural production for the household's own consumption. However, some questions of this block still remained in the household questionnaire. The household questionnaire was further supplemented by questions on household members who have moved abroad. Nevertheless, compared to the pilot census, the time gain in the actual census was at least 10% with those questionnaires as well. A notional full set, comprising the household questionnaire, the dwelling questionnaire and 2.3 personal questionnaires (based on the preliminary estimate of the size of an average household), was filled out in 51 minutes online and in 22 minutes in the interviews. Compared to the pilot census, the time spent on the questionnaires had decreased by 10% and 16%, respectively.

Figure 12. Time spent on filling out the household questionnaire, by the number of household members and method
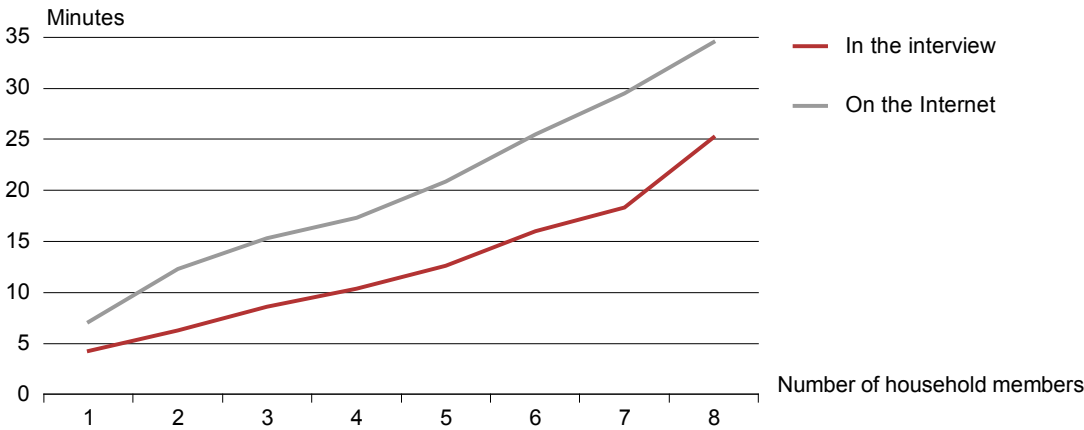
The time spent on filling out the household questionnaire was strongly dependent on the number of household members – a higher number of household members meant that more relationships had to be entered in the relationship matrix. The table of relationships had a relatively large number of checks: for instance, soft checks were associated with the age difference between household members (an unusually large age difference between partners, an unusually small age difference between grandparents and grandchildren) while strict checks blocked any impossible answers (e.g. father is younger than the son).

To some extent, the time required for filling out the household questionnaire was increased by the need to enter the data on temporary residents and close relatives who had moved abroad. However, their impact on the overall time requirement was not too strong, as they only constituted 6% of all persons enumerated.

The enumerators were able to work faster due to the experience gained through practice. This is illustrated by Figure 13, showing the change in the average time spent on filling out a household questionnaire over the period of the census interviews.

Figure 13. Average time spent on filling out a household questionnaire during the period of interviews
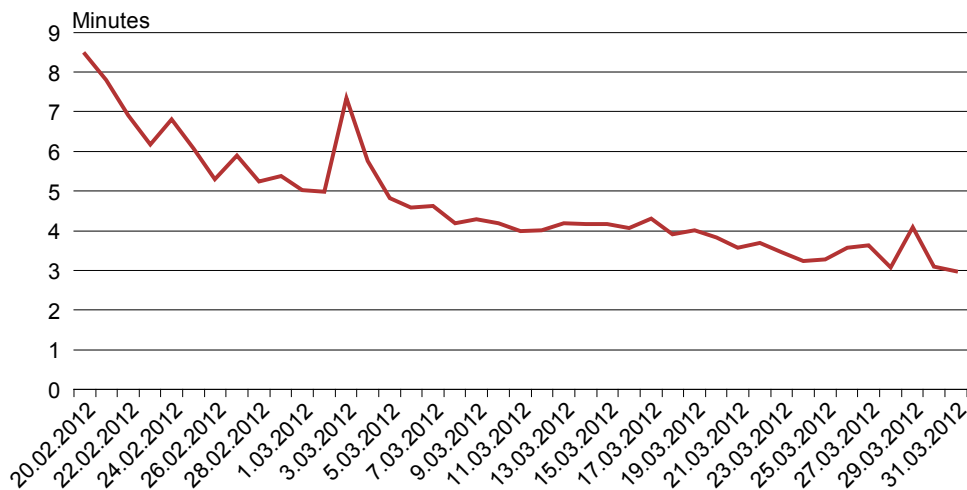
DATA PROCESSING

Preliminary data processing started immediately after the first personal questionnaires had been received. At the first stage, the task was to determine the persons and dwellings that had been enumerated, to avoid bothering people whose data had been submitted. The term for eliminating any duplicate personal questionnaires was relatively short, as it had been promised to announce the total population size within two months after the census.

Personal questionnaires received during the census

All persons who were subject to enumeration, i.e. the permanent residents of Estonia, had to fill out a personal questionnaire by answering 10–40 questions (excluding additional guidance questions). The number of questions could be even higher if the respondent knew a large number of foreign languages. The number of questions and answers was also above average for persons who had multiple places of residence, had lived abroad, or were women who had given birth. The lowest number of questions was required for small children – this explains why parents were able to fill out the children's questionnaires in such a short time.

In addition to permanent residents, data were also collected on other persons who were, according to the standard definition, not included among the persons subject to enumeration (Estonian population). On such persons, a so-called 'short personal questionnaire' was filled out by the same household member who filled out the household questionnaire. Each household was initially asked whether their dwelling included any temporary residents. If the answer was affirmative, there were some further questions about those persons, pertaining in particular to their place of permanent residence and identification data (name, personal identification code, sex, age, place of birth, citizenship). There were also questions about the connection between the temporary resident and the household, as well as the duration and reason for their stay in Estonia. Those temporary residents who had a place of permanent residence in Estonia were also subject to enumeration and, in many cases but not always, they were indeed enumerated at their place of permanent residence. The temporary residents who are permanent residents of a foreign country are not included among the permanent residents of Estonia, but some information concerning them is still traditionally provided in the census output. The enumeration of temporary residents was not mandatory in PHC 2011, but it was done as a matter of tradition. The enumeration of temporary residents also enables to collect data on those permanent residents who are relatively mobile. The questionnaires of temporary residents amounted to about 3% of all personal questionnaires received.

Each household was also asked about any members who had moved abroad between 2000 and 2011. In this case again, the questions were used to collect identifying information and some additional data, incl. the time of departure and the name of the country of current residence. The final number of the questionnaires on persons who had left Estonia also amounted to approximately 3% of all personal questionnaires. It was somewhat worrying that the percentage of household members marked as having left Estonia was lower in the e-census than during the census interviews (2% and over 4%, respectively). Even though it could be partially explained by the differences in the respective groups of respondents, it is more likely that some online respondents failed to mention their relatives abroad while in the interviews the enumerators could specifically point out this question, achieving a better coverage of that particular group.

Persons who tried to fill out their questionnaire abroad were first asked to specify the country of their permanent residence. If it was not Estonia, filling out the questionnaire was discontinued and the person was registered among those who have left Estonia. The number of such contacts was over 4,300.

In some cases, data on persons were collected with the help of lists. For instance, according to international agreements, the category of Estonian permanent residents includes Estonian residents who are on a diplomatic or military mission, irrespective of their current location. Some of them were unable to fill out the census questionnaire online and they were enumerated according to official lists, with some questions being left unanswered.

Persons living permanently in various institutions (care homes, custodial institutions, etc.) were generally enumerated using the usual method, but the corresponding institution was specified as their household (the so-called 'institutional households'). In exceptional cases, lists were also used to enumerate the members of institutional households – for instance, in the case of care home residents who were unable to answer the basic questions in the questionnaire and whose guardian could not be contacted. Lists were also used for assistance for the enumeration of the homeless.

Duplicate questionnaires and handling of duplicates

For various reasons, several personal questionnaires were received for some persons. As over-coverage is not a problem if people can be identified on the basis of the personal identification code, the methodology did not include any special measures to prevent duplicates. Instead, the presumption was that the more information we receive the better. As a result, there were many people with several personal questionnaires. Even though a person could not fill out two

personal questionnaires online (nor make corrections in a questionnaire after it was marked as completed), it was sometimes the case that a person's information was first entered by a household member (or even several household members) and then the person himself/herself filled out a personal questionnaire. For some children, personal questionnaires were filled out both by the father and the mother.

The submission of several personal questionnaires was expected in the case of those Estonian residents who had a temporary place of residence in addition to the place of permanent residence – they were enumerated at both places of residence. As the personal identification code makes it easy to identify and eliminate any redundant questionnaires filled out for a person, it was not a major issue in principle, even though it did somewhat complicate data processing.

The majority (about 90%) of the personal questionnaires received during the e-census included a personal identification code, which enabled continuous evaluation of the percentage of duplicates and monitoring of the actual number of persons enumerated. The share of duplicates among personal questionnaires was relatively low during that period. It is worth reminding that the actual personal identification codes, which would have enabled the identification of the persons, were never used in the analyses. Instead, the codes were converted to an encrypted form, preventing the possibility of linking the data with a particular person while retaining the possibility to link a person's data from various sources (census, registers, lists). The actual personal identification codes were only used by the identification code operators, who were responsible for correcting inaccurate identification data and adding personal identification codes from the Population Register to those questionnaires that did not include a personal identification code. They were not able to see the contents of the questionnaires. Most of this work was automated and human intervention was only required in the case of missing or conflicting data (e.g. missing personal identification code, different forms of the same name or different dates of birth in duplicate questionnaires or in a questionnaire and a register).

During the census interviews, nearly 40% of the personal questionnaires received did not include a personal identification code. It seems that some enumerators did not even ask this, because the questionnaire could also be filled out without the code. In some cases, the respondents did not want to disclose their personal identification codes to the enumerator, fearing that the enumerator could misuse this information, e.g. by taking an SMS loan. However, there were no reports of misuse of personal data in the census. As many personal questionnaires did not include a personal identification code, it was initially impossible to estimate the share of duplicates among the incoming questionnaires, but the share did increase for several reasons, especially during the second stage of the census interviews. By the end of the e-census, duplicates constituted about 3% of the questionnaires, while by the end of the census interviews, the share of duplicates had risen to 10% of all questionnaires. An increase in the number of duplicates was predicted, but the actual increase was somewhat bigger than expected. This was caused, firstly, by the second stage of the census when some people were enumerated for a second time and, secondly, by the use of lists and other sources for finding persons subject to enumeration. While the relatively large number of duplicates did increase the workload on the operators, it also improved data quality since, in several cases, the comparison of duplicates resulted in a greater number of responses to individual questions. According to the developed priority rules, the duplicate questionnaires were used to maximise the amount of reliable information collected on the respondents. The primary questionnaire chosen from among the duplicates was called the **original**. The information of the original was supplemented as necessary and possible with the information from the other duplicates.

The number of duplicates was higher (as expected) in the case of those people who had moved abroad, because the information about them was often provided by several relatives – parents, children, siblings, or the person himself/herself trying to fill out the questionnaire in a foreign country. However, even a collection of all this data does not provide a full overview about the people who have left Estonia in the past twelve years, because some of them could no longer have close relatives still living in Estonia and it is also unclear whether all respondents considered it necessary to provide information on the people who are now abroad.

The duplicates were processed according to the following priority rules:

The **first priority tier** was based on the person who filled out the questionnaire

- Questionnaires filled out in person or by the parents of a minor child;
- Questionnaires filled out by another household member;
- Somebody who started to fill out the questionnaire (logged in) in a foreign country was generally registered as having left Estonia.

The **second priority tier** was based on the person's relation to the household (permanent resident, temporary resident, has left the country)
- If two persons were ranked the same in the first category of priority, preference was given to a permanent resident;
- A temporary resident and a person who had left the country had the same level of priority.

The **third priority tier** (applied when the first two tiers did not produce the needed data) was the method of interviewing
- Computer-assisted personal interviewing (CAPI) (pen-and-paper interviewing, i.e. PAPI, in exceptional cases) and the combined method (CAWI + CAPI) had priority over computer-aided web interviewing (CAWI).

The **fourth priority tier** was the time of the interview
- In the case of CAPI, PAPI and combined methods, preference was given to later times;
- In the case of CAWI, the earlier time was preferred.

Basic principles of address specification

Address specification is important in the case of permanent residents and was based on the described priorities. The address used in the census was the address of the original. Address specification was also required if a temporary resident was re-categorised as a permanent resident. In this case, the following priorities were applied:

- Full address of the place of permanent residence if included in any of the duplicates;
- Address of the place of temporary residence where the person is listed as a household member (if there were multiple addresses, preference was based on the rules concerning the interviewing method and time);
- Address according to the population register if no address submitted in the PHC could be linked to a building.

Determination of the household of a person with a specified address

This relationship was defined by the household of the dwelling if the person was listed as a member of one of the households in that dwelling. If the person did not belong to any households in that dwelling, he or she was registered as a separate household.

Who were enumerated as permanent residents of Estonia?

The determination of permanent residence in Estonia was generally based on international requirements. The only problems occurred with persons who had several personal questionnaires with conflicting information. Certain rules were formulated to resolve such situations. The following persons were considered permanent residents of Estonia:

- Persons who had a complete personal questionnaire filled out;
- Persons listed as members of a household of permanent residents, with no other information available;
- Persons with a questionnaire of a temporary resident, specifying that the person's place of permanent residence was in Estonia;
- Persons whose residence in Estonia was verified according to various lists.

Who were registered as having left Estonia?

A new category of persons that was the subject of this census included former permanent residents of Estonia who had left the country (in the past 12 years). As population censuses are generally not well-suited for measuring emigration, there is no internationally recognised methodology for that purpose. However, some principles can be derived from the methodology of measuring immigration. A key consideration is that the person, who has left the country, cannot provide information about his or her own emigration during the census. This usually also applies to the family members of such persons, as persons residing in different countries cannot constitute a household. Consequently, information can only be provided by a more distant person, who could even be a close relative but whose connection with the person or knowledge about the actions or intentions of the person cannot be verified.

The concept of emigration is closely linked with the time after which a person is deemed to have the status of a permanent resident of another country. In the case of immigrants, this period is set at 12 months and, by way of derogation, the status of a permanent resident can also be registered if a person has not stayed in a country for 12 months but states that he or she intends to do so. This criterion cannot be applied to emigrants, because they or their household members cannot be asked the question about their intended length of stay in the foreign country, and we have to make do with indirect information.

In the PHC 2011, all persons who had left Estonia before 2011 according to their relatives and on whose residence in Estonia no information was received during the census were enumerated as having left the country. Furthermore, persons studying at foreign universities or post-secondary vocational education institutions for a period exceeding one year were also deemed to be residents of a foreign country (based on international agreements). The category of persons residing abroad also included all temporary residents whose place of permanent residence was located in a foreign country (some of them had left Estonia in the past twelve years), as well as the persons who tried to fill out the census questionnaire abroad and whose enumeration was discontinued as they did not have a place of permanent residence in Estonia.

After all these relatively clearly specified categories of persons had been processed, there remained some others in whose case it was rather complicated to make a decision. They numbered less than 5,000 in total and included:

- Persons with conflicting data – there were one or several questionnaires in which they were listed as having left the country and one or several questionnaires in which they were entered as temporary residents who had a place of permanent residence in Estonia;
- Non-students who had left Estonia in 2011 and had not spent enough time in a foreign country to gain the status of a resident in that country, and from whom it was also not possible to ask about their intended length of stay in the foreign country.

Additional data sources were used to determine the status of such persons. The persons who had moved in 2011 were registered as having left Estonia if they had changed their place of residence and the population register indicated that they had a place of permanent residence in a foreign country. In the remaining cases, the decisions were based on any comments included in the census questionnaires and/or the persons' activity in Estonian registers.

How many are we?

The total size of the population is always the first question when it comes to census data. To answer this question, it is necessary to:

- take into account all personal questionnaires received;
- remove all duplicate questionnaires;
- make a decision on persons who simultaneously have some attributes of a permanent resident and some attributes of a non-resident.

Providing information on the geographical distribution of the population also requires determination of the place of permanent residence of all persons. This stage was much more difficult during this census than it has been before when the enumerators enumerated persons by residential buildings and dwellings, always entering the address. The preliminary number of permanent residents was announced in May 2012, less than two months after the end of the census. Subsequently, after the completion of data processing and additional verification, the total population size was corrected by only a few hundred persons (the number increased). The final population size according to the census was 1,294,455 persons.

However, due to under-coverage, this population figure was somewhat below the actual population size. The calculations based on data from registers (Tiit 2012) indicated that the rate of under-coverage was about 2.1–2.2% and, consequently, the accurate (estimated) population figure was a little over 1,320,000. Nevertheless, all census results were presented using the census population, i.e. the population figure of 1,294,455 persons. The revised population number will be used in subsequent population statistics, but the census results were presented based on the number of enumerated persons. The population number used in population statistics was fixed with a news release published at the beginning of 2014, by which there were 1,325,217 permanent residents in Estonia as at 1 January 2012, meaning that under-coverage amounted to 2.3%.

Data processing after the census interviews

Data processing continued after the interviews when all questionnaires had been received. A characteristic feature of censuses, like any other surveys, is the fact that the results cannot be presented immediately after the end of data collection – we also need a stage of data organisation, which is relatively difficult to grasp for the outsiders. While data entry (from paper questionnaires to computers) took up a large portion of that stage in the previous censuses (from 1922 to 2000), data organisation is now mainly reduced to data verification and the encoding of a rather small number of text elements. Nevertheless, intense work on data organisation, carried out by trained operators, continued after the end of the interviewing period of the 2011 census. This time, a part of the data organisation process was automated. For that purpose, a systemic tool – an operator's desktop – had been developed, providing convenient access to the information required for the task at hand. The main tasks included the following:

- Determining personal identification codes based on personal data (in the case of personal questionnaires without a personal identification code);
- Ordering and standardisation of addresses;
- Encoding of textual responses.

A special software application had been developed for that purpose, including specific desktops for ID code operators, address operators and encoders. This was also a means to ensure data security – only persons in a particular role could use the respective desktop. The data accessible on each desktop was limited to the extent needed for the performance

of a particular task. At the same time, the system enabled the automated completion of simpler and standardised operations.

The unexpectedly high participation rate in the e-census also affected the speed of data processing: during the period of the e-census, the system resources were used for the census and the capacity available for data processing was below the actual requirements.

Personal identification code operators

The ID code operators were responsible for the **identification of persons**, adding **personal identification codes** to the data sets with a missing code, as well as checking the match between the available personal identification codes and names.

Persons were identified automatically if they personally logged in to the system to fill out the questionnaires in the e-census or if the name they entered was an exact match for the name in a register or if they were on the enumeration list of the respective enumerator during the census interviews. There were ten personal identification code operators in total and they worked from the first day of the census until the middle of April, i.e. for 4.5 months. The operators solved over 175,000 cases related to personal identification codes, with further specification or identification required for every tenth person enumerated online and for more than a third of the interviewed persons. The personal identification code cases were resolved rather successfully, with less than 1,500 enumerated persons remaining without a personal identification code. The number of permanent residents among them was less than 400 (*ca* 0.03%).

Figure 14. Tasks of personal identification code operators by week
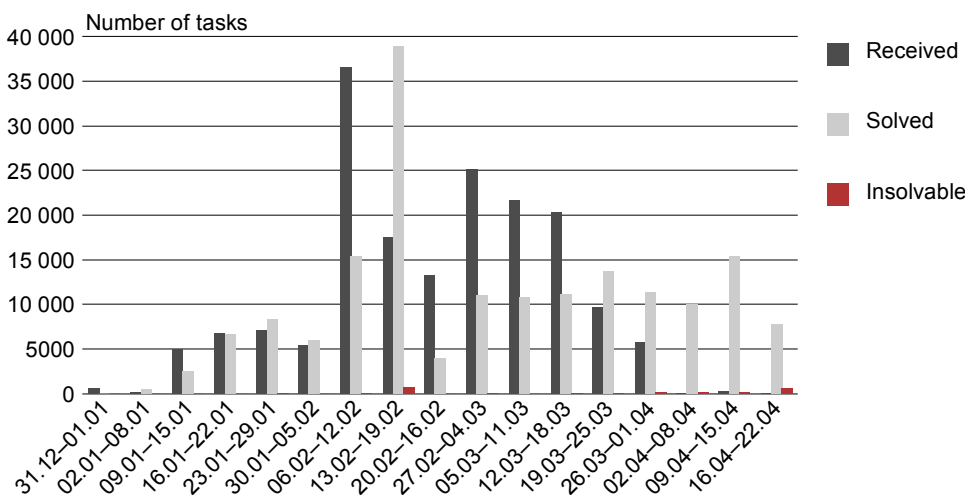
Address operators

The **organisation of addresses** (performed on the desktop of address operators) turned out to be a rather time-consuming and complicated task, requiring significantly more time than initially planned. The address operators had to identify the addresses of permanent and temporary places of residence, convert any fully or partially free-text addresses into the standard ADS format, and link the addresses with building IDs. While several types of mistakes were made when entering the addresses, they usually shared a common trait – the new address standard had not been fully adopted and accepted yet. People in villages are not accustomed to entering the address details beyond their mailbox number or village name. Even in villages where streets or farms have been named and numbers have been assigned to houses and apartments, people still usually only write their name on postal consignments – at best adding something about their house – and this is quite different from the address format used for those houses in ADS. There where instances when people entered the former land unit numbers from the first independence period of the Republic of Estonia or house numbers that were used in village centres in the 1950s. Sometimes people did not specify the actual location of their house on the map, marking the location roughly somewhere in the village, e.g. near the church.

It was also discovered that it was indeed difficult to apply the standard to certain types of addresses, e.g. residential buildings consisting of several parts (blocks). Difficulties were also created by location and street names which could be written in different ways (e.g. streets named after persons where the first name of the person should be included in the street name). The six address operators worked for four months, resolving over 120,000 cases (i.e. one-fifth of all dwelling addresses had issues that needed resolving). Despite the number of questionnaires filled out online being

almost twice as high as the number of questionnaires from census interviews, it was the latter category that created more work for the address operators. Nearly a quarter of the cases remained unsolved by the end of the planned work period and work on them had to be continued for several months in parallel with other data organisation tasks.

Figure 15. Tasks of address operators by week
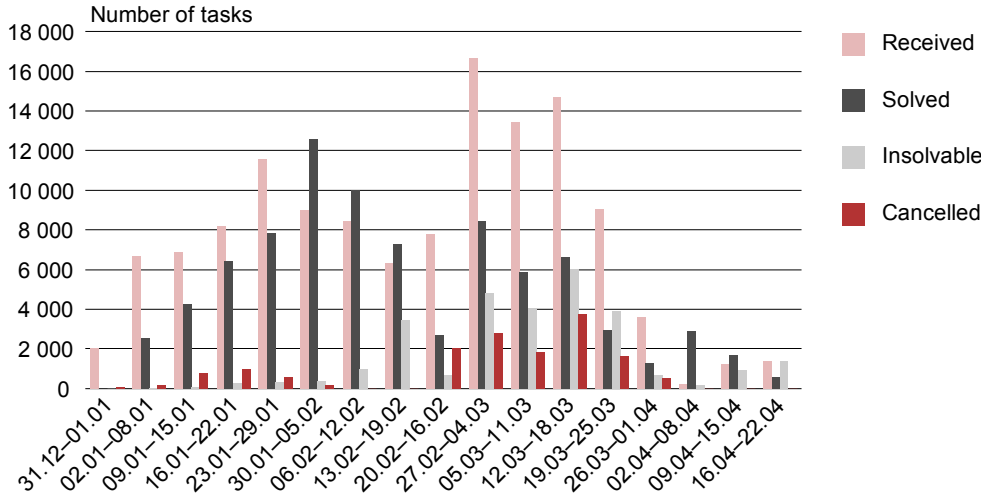
Encoding

Encoding was the most labour-intensive part in the data organisation and processing stage. In general, encoding means the division of individual objects into groups (types, classes) based on particular attributes, so that each object would only belong into one group. Such grouping is based on classification. The system attempted automated encoding of any free-text answers, or strings, in the questionnaires of the PHC 2011. This was done by checking whether a task-related dictionary included an exact match for the string. The encoders had to manually encode any strings which did not have an automated match. The encoders had to also consider any comments added to the questionnaires and other related attributes.

Full encoding was required in the case of employment attributes 'area of activity' and 'occupation', which had to be presented in accordance with international classifications (NACE and ISCO). Partial encoding was required for a number of attributes where questions were answered in three stages, with a free-text field being the third stage. This included all country-related attributes (more than ten in total, incl. country of birth, country of employment, etc.), ethnicity and citizenship, native language, foreign languages, dialects, and religion. Even though the number of answers that required manual encoding was relatively low with regard to those attributes, the encoding itself was quite complicated as the respondents themselves had not been able to make a decision.

The dictionaries used for encoding in the PHC 2011 were based on international classifications or lists created specifically for PHC 2011.

In addition to classification entries (originals), the dictionaries also include synonyms (alternative place names, incorrectly written words, names in dialects, etc.) and entries not used in automated encoding (the so-called 'banned words' – designations that are too general and uninformative for effective classification, e.g. 'manager'). The same dictionary was used for both automated and manual encoding. The codes of missing values were the same for all attributes.

In addition to dictionaries, several other **help resources** were used in encoding: the operator's manual, all classification and dictionary files, other training and work files in a restricted-access folder, print versions of the economic activity and occupation classifications. The statistical register of economic units was a very helpful resource in encoding economic activities, while the maps and map files created by GIS experts were extremely useful in resolving any questions associated with Estonia's former administrative divisions. Several issues were also resolved using web search.

Contracts had been signed with **outside experts** who helped to resolve more complicated cases. Consultations were provided by Mihkel Reispass from Statistics Estonia on occupations and economic activities, by Peeter Päll from the Institute of the Estonian Language on countries and citizenships, by Jüri Viikberg from the Institute of the Estonian Language on languages, ethnicities and Estonian dialects, and by Ringo Ringvee from the Ministry of the Interior on religion. The more difficult cases were collected in thematic groups and sent to outside experts once every one or two weeks.

Table 5. Encoding tasks of the population census by dictionary

Stage 1 (25.04.12)	Encoded	Manual encoding	Manual encoding, %	Automated encoding	Automated encoding, %
VS-AM	646 393	634 101	98,1	12 292	1.9
VS-EMTAK	870 683	799 745	91,9	70 938	8.1
VS-US	2 750	1 871	68,0	879	32.0
VS-RTK	13 764	6 938	50,4	6 826	49.6
VS-KOD	837	598	71,4	239	28.6
VS-KEEL	1 664 713	6 573	0,4	1 658 140	99.6
VS-RA	1 633	1 197	73,3	436	26.7
VS-MR	10 161	5 374	52,9	4 787	47.1
VS-EHAK	84 887	84 887	100,0	0	0.0
Total	**3 295 821**	**1 541 284**	**46,8**	**1 754 537**	**53.2**

The high number of tasks associated with the VS-KEEL dictionary was due to the fact that the encoder's desktop also included the answers on such foreign languages which were selected from a pre-defined list while filling out the questionnaire. The percentage of automatically resolved language tasks is very high for the same reason. Curiously, many respondents had also indicated knowledge of programming languages when answering the question on language skills. While the command of a programming language is certainly a valuable skill, it was not a suitable answer for the PHC question on language skills, as the focus was on the command of natural languages.

Automated encoding was not used in the tasks associated with the VS-EHAK dictionary, as they generally required the use of several response fields. The encoders were also unable to complete their work within the envisaged timeframe, with 13,000 encoding tasks or 0.4% of the total number of tasks remaining unresolved by the due date. However, all the tasks were resolved in the end.

Figure 16. Tasks of encoders by week
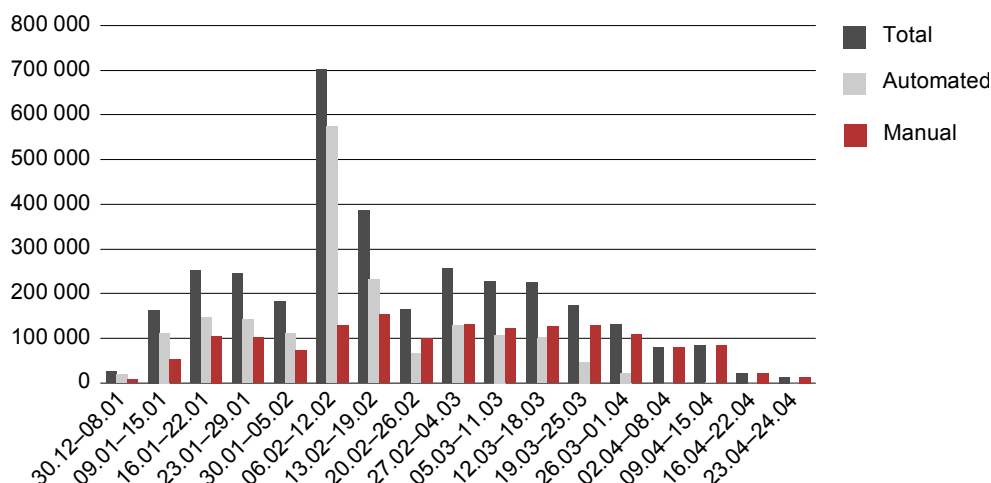
Filling data gaps and eliminating incorrect values

The identification of objects and the encoding of textual values are not the only jobs required for organising a data set and preparing it for analysis. Some values are almost always missing from questionnaires – even if failure to answer is made relatively difficult for respondents by the use of checks. There are also some values that are probably incorrect (again, despite the checks), sometimes caused by technical faults. Even though the data were checked in real time as they were entered in the PHC 2011, additional (systematic) checks were used in data processing to detect any conflicts between different attributes. Such clearly conflicting values were removed, but the overall number of such values was rather insignificant.

However, data gaps can cause a significant loss of information. For instance, if an average of 1% of the values of an attribute is missing, a table with six attributes could include 6% of objects with incomplete information, and this would be a significant loss. To avoid this, attempts are always made in surveys to fill any potential data gaps if possible. There are several options for filling data gaps:

- Imputation – data gaps are filled according to mathematical rules, using the values of other objects in the same data set. Various algorithms have been developed for this and the respective procedures are very common;
- Data gaps are filled on the basis of other values in the same questionnaire, using logical and statistical procedures;
- Data gaps are filled by using data from another source.

As the methodology of the Estonian Population and Housing Census 2011 prescribed the use of registers and the registers had been extensively tested during census preparations, the third option was chosen for filling the data gaps in PHC 2011 – missing data were replaced with values derived from registers. The preparatory work had been done before the census, by preparing the data for pre-filling the answers. The Population Register, the PHC 2000 database and the Estonian Education Information System were the main sources used to fill the gaps in personal questionnaires, while the Register of Construction Works was used for dwelling questionnaires. These registers did not enable filling all data gaps and, consequently, a considerable number of missing values remained in the data set in the case of some attributes (1.5% on average) (see Tables 2 and 3 in Annex 1). Imputation was not used in the organisation of the data of PHC 2011 – not for census objects nor for individual attributes.

OUTPUT COMPILATION

It would be erroneous to think that output tables can be prepared right after the end of data processing, the identification of all the items subject to enumeration, the correction of mistakes and filling the gaps in the collected data. It is possible in the case of simple characteristics like sex, age, or legal marital status, but in most cases there is one more step before final output.

Output describing persons

The typical feature of the Population and Housing Census 2011 is that several output indicators are not direct responses to the questions, but instead consist of several answers (sometimes based on other sources), combined by using special rules (algorithms). Detailed information about forming new characteristics is available in Table 1 of Annex 1 (p. 64), which shows the census characteristics and other sources used for their compilation.

Educational attainment. The census plan prescribed using data from the Estonian Education Information System (EHIS) to determine those attending school at the moment of census, as well as their level of education and the location of the educational institution. Although EHIS provides a good overview of educational attainment in Estonia, it does not cover studies abroad (distance learning). However, the shortcoming was not remarkable, because persons who were abroad due to tertiary or post-secondary vocational studies (at least for a year) were not enumerated as Estonian residents, and young people who were abroad due to upper secondary studies had to be enumerated at their parents' place of residence. EHIS also provided reliable data for filling the gaps and finding out possible discrepancies.

Obtained education. It was rather complicated to present the educational data on the international scale (ISCED 97), which is the classification required by Eurostat. On that scale, the description of the levels of education prescribes using both the characteristics indicating the type of school attended and the characteristics indicating the previous education required, as well as whether the school provided general education besides a vocation/profession. A detailed education scale was also necessary to compare the population's level of education to the previous census. For that purpose, the comparison involved a scale with 14 levels. Although one question concerned literacy, it was not included in the published tables – the share of illiterate persons in the total population was insignificant (mostly people with severe disability). In order to create categories suitable for output tables, algorithms consisting of several dozens of rows (conditions/references) had to be used.

Employment. The employment of a person was determined based on three questions developed by the International Labour Organisation (ILO). Unemployed persons included only those persons without work at the moment of census who were not temporarily absent from work, who were actively seeking work or preparing for starting a business, or were ready to start working within two weeks if given the opportunity. Employment constitutes a part of the socioeconomic status, which combines the information collected about the economically active (including the unemployed) and the economically inactive (persons not working due to other reasons). Additionally, new characteristics were established for the unemployed – **economic activity of the last job** and **occupation at the last job.**

When preparing the questionnaire, the preparation team avoided questions with obvious answers. Persons at the age of 15 or younger were not asked about their religion, education, marital status, number of live-born children, or employment. The education of that age group (less than primary education, primary or lower secondary education) was determined by their educational attainment (from EHIS); school-aged children cannot have an employment relationship (as the main activity). The **age at the birth of the first child** was not asked about because there is sufficient information available about that in the statistical births database.

The values of some characteristics were combined with responses to other characteristics. The source of subsistence was not asked from persons aged 15 or younger – it was presumed that those living in a private household would indicate 'maintained by other persons' and those living in an institution would indicate 'maintained by institution'. The questions about living abroad and the previous place of residence were not asked from persons who affirmed that they had lived their entire life in the same place.

Command of foreign languages, which in this form (for persons at least three years of age, up to 20 foreign languages besides mother tongue could be indicated) was included in the population census for the first time, required defining several additional characteristics, the majority of which were not presented in a table format, but were used to create output. In the case of the command of foreign languages, a list of all languages with at least five speakers was published. The distribution of spoken languages was also ascertained (in tables, up to 8 spoken languages are indicated), and several other characteristics were shown with five most common foreign languages in Estonia.

The characteristics of mother tongue and the command of foreign languages were combined into the indicator[a] **command of Estonian language** – speaks Estonian as a mother tongue, speaks Estonian as a foreign language, does not speak Estonian.

For the first time, information about health-related limitations (limitations of everyday activities due to health problems) was asked about and used in calculating **disability-free life expectancy**, which was also done for the first time based on census data. The estimation of this characteristic required using mortality indicators (received from regular mortality statistics), the tables included both disability-free life expectancy and **life expectancy**.

Native and foreign-origin population. The information on the birthplace of persons, their parents and grandparents was used to create the four-level characteristic of **native and foreign-origin population**, which is an extension from the three-level concept of origin used in previous studies. The levels are defined as follows:

- native population – persons permanently living in Estonia, at least one of the parents and at least one of the grandparents of whom were born in Estonia.
- first generation of foreign-origin population – persons permanently living in Estonia who and whose parents were born abroad;
- second generation of foreign-origin population – persons permanently living in Estonia who were born in Estonia but whose parents were born abroad;
- third generation of foreign-origin population – persons permanently living in Estonia of whose parents at least one was born in Estonia but whose all grandparents were born abroad.

Some of the personal characteristics, however, are found on the basis of the information provided in the Dwelling and Household Questionnaire, for example, a person's status in the household and a person's status in the family.

A series of questions was also asked about persons who are not permanent residents of Estonia – i.e. temporary residents and the emigrated close relatives of household members. Temporary residents were asked about their sex, date of birth, place of residence, country of birth, citizenship, and the duration of their stay in Estonia. Those emigrated from Estonia were asked about their sex, date of birth, country of residence and time of emigration.

Output describing dwellings and residential buildings

A **dwelling** is one of the main objects of the population and housing census. In most cases, a dwelling is a **conventional dwelling** – a one-family dwelling, a box of a terraced or semi-detached house or apartment, if it is suitable for all-year-round habitation. Another type of dwelling is an **accommodation room**, e.g. a room in a dormitory. In addition to these, any other type of housing was regarded as a dwelling, if at the census moment at least one person was residing there permanently, despite being unsuitable for a conventional dwelling, e.g. a summer cottage. Such a dwelling is (paradoxically) called a **non-dwelling**. Some of the persons enumerated – members of the so-called institutional households – were also living in institutions (e.g. children's, youth and care homes, penal institutions, monasteries, etc.), but the number of rooms of institutions is not recorded in the population and housing census.

As noted, all the enumerated dwellings were identified by an eight-component address in the ADS-system, the first three components of which contained the EHAK (Classification of Estonian Administrative Units and Settlements) codes – county, local government unit (city, rural municipality, city district) and settlement unit (city without municipal status, town, small town, village) codes respectively. The next levels specified the address by the name of a small place, street, farm or building, the number of the building and apartment (part of the building). In addition to that, the location of each building was marked on a map. Dwellings are identified by an unequivocal index present in all dwelling-related registers.

However, there were some persons (including those temporarily abroad, the homeless, etc.) for whom the exact address and location of the place of residence remained unknown. The number of these persons totalled *ca* 4,000. A fictitious place of residence was created for them, located in the centre of the settlement unit (the so-called centroid). Naturally, no other features of a dwelling were determined for such a "place of residence".

If the location of the dwelling was known, it was possible to determine every enumerated dwelling (and also its residents) in a suitable square on the grid map of Estonia (a map divided into grids with a side corresponding to either 1 km, 500 m or 100 m), enumerate the number of dwellings and persons in a given grid and get an overview of the change in population density on the territory of Estonia. Such linking of persons enumerated and the location of the dwelling is also necessary to determine localities, which do not necessarily coincide with settlements, especially in urban sprawl areas, and to supplement the personal and dwelling characteristics with the size of a relevant locality (number of residents). The size of a locality (on a pre-determined scale) is a core characteristic (see p. 16).

---

[a] The description of all characteristics generally omitted the standard divisions "unknown" and "total".

Residential building. A dwelling is an enumerated item, whereas a residential building (a building) is derived on the basis of dwelling data. In the case of one-family dwellings (private houses), the residential building and dwelling coincide. As for apartments or parts of a building as dwellings, there are usually several of them in one residential building. A dwelling is characterised by an address and a dwelling index (specified in the course of data processing), residential buildings include all buildings containing at least one conventional dwelling. A residential building is described by two characteristics – the type of residential building and the number of rooms in the residential building. These are derived by linking the residential building to the dwellings it contains. The said characteristics are not calculated in the case of fictitious residential buildings (centroids).

Dwelling characteristics. When speaking of the **occupancy of a dwelling** one has to take into account that the residents include only permanent residents of Estonia – if the dwelling accommodates, for instance, diplomatic staff of a foreign country or temporary residents, then it counts as not inhabited. The **number of permanent residents in a dwelling** is calculated for all dwellings by adding up all the members who are the usual residents of the household permanently living in the dwelling. This figure is also calculated for centroids. If there are no households that permanently live in the dwelling, the number of usual residents of that dwelling is zero.

The number of permanent residents is used to calculate indicators expressing population density, such as the **number of rooms in the dwelling** and the **area of the dwelling per resident**. All these, as well as the indicators of the technical equipment of the dwelling, are subject to publication and most of them are core characteristics, but in the case of European output, the rooms include the kitchen. The only question added to the core questions regarding dwellings was the existence of a sauna, which was asked about from those households that did not have any other washing facilities.

Output describing households and families

Households are divided into private households, institutional households and households of homeless persons. An institutional household consists of the residents of an institution (care home, prison, monastery). Homeless persons are persons who did not have a place of residence at the census moment. A private household consists of persons who live together (in a common conventional dwelling) and share the available household facilities (common budget and food). A person living alone is also a household.

A **household** is a census unit for which the list of members and their mutual relationships are established in the census.

A **family**, on the other hand, is a derived concept. A **family** or a **family nucleus** is either:

- a legally married couple or a cohabiting couple without children;
- a legally married couple or a cohabiting couple with a child or children;
- a single parent with a child or children.

A family is always part of a household[a], i.e. a family must have a common dwelling and a shared household. A household may contain either one or more families, but there are also households without any families – these are all one-member households. If a household consists of one or many families, it is necessary to find out its/their members, taking into account that a person can only belong to one family and the younger family takes priority over the other one. However, in most cases, families and households tend to coincide.

For each household member in a family (i.e. family member), his or her **status in the family** (partner, parent or child) is determined. The members who belong to the family determine the **type of family**. The main types of family include a married couple with or without children, a cohabiting couple with or without children, a single mother with a child or children, a single father with a child or children. The number of types increases when differentiating the family by the age and number of children or by the sex of the cohabiting partners (same-sex or opposite sex). Here, it has to be kept in mind that there are no age-related restrictions with regard to a child as a family member – even a 60-year-old child can belong to the same family with his or her parent(s), but the important thing is that he or she has no child, spouse or partner (in this family).

The **household structure** arises from the structure of families included in a household. According to the number of families per household, there are non-family households, one-family households and multi-family households. A household may also contain persons who do not belong to any of the families in the household, i.e. they are other persons (extra members). Thus, besides the types defined for families (e.g. married couple with children) there are other types for a household if it has extra members (e.g. married couple with children and other members). Non-family households include households with one or several members, e.g. siblings living together. If a household consists of a grandparent and a child or children, but no parents, then it is referred to as a **household with a missing generation.** The household structure also provides for various unusual household types such as a household with same-sex partners (for the first time in Estonia based on the PHC 2011 data). Another type of household that is increasingly frequent in

---

[a] The census covered only families in private households because the number of both institutional and homeless families is vanishingly low (if any exist at all).

Estonia is a **reconstituted family** (one or both partners have formerly constituted a household in partnership with some other person), and its special subcategory **patchwork family** (at least one child is not a common child of the partners).

After determining the type of household, it is possible to ascertain a **person's status in the household** for each household member – it is always determined for each person. The status in the family, however, is not determined for the persons who do not belong to the family nucleus. Still, a household often stands for family, and the status of all household members coincides with their status in the family.

In order to analyse both households and families, indicators were developed to describe the households/families by the members' age and status, as well as the number of members at a different age (e.g. **age and number of children**, **age of the youngest child**, **number of members and dependants aged 18 and younger, number of employed members**), in the case of households also by the number of generations (**number of generations in household**). These indicators are used in the output intended for Estonian consumers and have not been included in Table 1 of Annex 1 (p. 64).

One table was published with regard to **institutional households**, indicating the type of institution, the person's age, sex, and nationality. In all output, the term household generally stands for private households.

PRESENTATION AND PUBLICATION OF CENSUS RESULTS

Output tables

Traditionally, the primary census output consisted in tables, which were printed out until recently. Already the first population census carried out on the Estonian territory (in 1881) resulted in a series of tables and analysis thereof, illustrated by colour-printed (*sic!*) images. Several sets of tables (one, in some cases two per each county) were also issued after the census of 1922. The situation was different in the Soviet times – census data remained unpublished, data tables were only available to specialists (for so-called official use only). The tables contained only numbers (mainframe printouts) and had no comments or analysis. For Estonia, the census tables for the 1959, 1970, 1979 and 1989 censuses were published in 1996–1997.

The Population and Housing Census 2000 produced 12 volumes of hard-copy materials published in 2001–2004. At the same time, the census team compiled electronic tables for publication in the Statistical Database. In the case of electronic tables, the consumer can select the required characteristic from three to four or sometimes even more characteristics, and generate a table in a suitable format for making additional analyses. The last electronic tables containing the data of PHC 2000 were published in May 2002, but some of them were supplemented in 2003. The Statistical Database contains 183 tables for the 2000 Population and Housing Census, which cover the entire spectrum of characteristics measured in the census.

The 2011 Population and Housing Census was planned so that remarkably less paper would be used than in the previous census. Besides the virtually paperless enumeration process, the team also planned using mostly electronic output.

Although initially one of the goals was to ensure the comparability of data from PHC 2011 and PHC 2000, there are still remarkable discrepancies between the tables of PHC 2011 and those of the previous census.

- PHC 2011 considers the needs of regional consumers – the majority of the tables provide information on the county level, several also on a lower level of settlement hierarchy. However, due to a significant decrease in the number of local government units compared to 2000 arising from changes in administrative arrangement, there is no great difference in the number of single cells (more than 22 million in both cases) in the output of PHC 2000 and PHC 2011.
- PHC 2011 tables generally contain more characteristics, which makes them more flexible to use.
- the output of PHC 2011 includes more tables (312 instead of 183)
- the presentation of data for PHC 2011 is somewhat more thoughtful and clear.

Various scales (aggregations)

For several characteristics, different output tables use different scales. For example, there is one age-related table, which contains 101 age values, ranging from 0 to 99 and 100+. In many cases, a scale with an interval of five years has been used (0–4, …, 80–84, 85+). In case a table contains multiple characteristics, age scales with a three- (0–17, 18–64, 65+) or five-year span (0–14, 15–29, 30–49, 50–64, 65+) have been used. A similar situation occurs for various characteristics with an initial high value count.

For example, the characteristic of place of residence/location has 5,262 values on the settlement level, 308 on the rural municipality level, and 19 and 23 values on the county level – in most cases, the table also includes relevant values for the interim count (rural settlements, urban settlements), the table of counties includes data about three major cities in Estonia as well. Besides these scales, scales limited by the number of population are used as well, e.g. data are presented for all settlements with at least 200 residents (less than 1,000 values) or 2,000 residents (163 values).

In the case of all characteristics with a large scale (countries, ethnic nationalities, citizenship, languages, dialects, religion, occupations), at least one table shows the full list of values (as detailed as the confidentiality requirements allow) and most of the other tables show a suitably limited number of values. For instance, the maximum number of values is 230 for occupations and 471 for economic activities.

Multiple scales are also used for many remaining characteristics (e.g. two scales for education – a more detailed one with 16 levels and a less detailed one with 7 levels). This is due to the fact that in tables with a large number of characteristics, the number of value combinations would reach millions when using detailed scales and most cells would inevitably remain empty. Using such tables can be inconvenient and time-consuming and would cause confidentiality issues. Therefore, most tables use less-detailed scales for the so-called background characteristics. The same principle has been applied to the output regarding core characteristics – there, too, all characteristics have several scales and the definition of hypercubes indicates which one to use.

Confidentiality check and hypercubes

The output of Census 2011 contains another feature not present in former census data. Namely, some tables contain intentional errors – small ones, but still noticeable to an attentive observer.

Today, it is necessary to ensure the protection of personal information. Nobody has the right to find out about anyone's health, dwelling ownership, biological or adopted children, etc. During the census, data is collected even in the smallest villages and groups of people, and thus, some tables could indicate that there is a man aged 40–44 in village N, and anyone who knows that man could find out about everything asked from that man. In order to avoid such a situation, it is internationally prohibited to publish tables that contain cells with values (frequency) of 1 or 2.

There are various methods to prevent such a situation, and they all alter the data set. One option is to leave the low-frequency cells unpublished, replacing their values with a symbol. Unfortunately, this does not always help – in most cases, a simple calculation is enough to find out the contents of the hidden cell by using the values of other cells. Another option is to add a random error to the results. Estonia chose the third option – in the case of a table with confidential values, all frequencies were rounded to base three. This removed all values '1' and '2' from the tables, and all the published numbers divide by three. The result is somewhat erroneous, but generally the error is relatively small.

Hypercubes. The Statistical Office of the European Communities (Eurostat) established very precise requirements for the output of Census 2011. The mandatory output consists of 60 tables with an average of 6 to 7 dimensions – the so-called hypercubes – which should allow a very detailed comparison of the EU countries. Still, there are many problems associated with hypercubes: on the one hand, these tables contain a great number of cells due to a multitude of dimensions, resulting in empty or almost empty cells even in large countries, and on the other hand, several cells contain figures, the publication of which is prohibited by confidentiality rules.

Publicity for the census

Publicity activities were crucial for ensuring the success of the census. Publicity events commenced early, already before the pilot census, and employed various media channels.

In Estonia, censuses have always been media events to some extent. Prior to a census, sometimes for several years, newspapers have discussed population problems, informed readers about the census aims, given instructions on how to behave during the census, and erased census-related fears. These media fragments constitute the majority of the information available telling us about the nuances and daily issues of previous censuses, as – unlike numeric data – there is not much text information left for several censuses. The involved parties have described and recorded only the preparations for the censuses of 1881 and 1922.

"The success of your census depends on how well you can sell it", said Pekka Myrskylä, a Finnish census expert at a seminar held in 2009. The Estonian experience confirmed his statement.

How to ensure a positive attitude of the public and media? The media was rather critical during the previous census in 2000. An article called "Poor people or poor census?" („Kas nädi rahvas või nädi loendus?") was published in the daily newspaper "Päevaleht" already a few days after the census had ended. The article cast suspicion on the census results and described the enumerators' mistakes regarding enumeration ethics and rules. In several countries, the public and media attitude was rather reserved and critical during this census as well. Naturally, this had its impact on the census results, too.

The information campaign for PHC 2011 was a professional one. It started already before the pilot census and followed a steady upward trend. The campaign slogan "Everyone counts" was characterised by openness and active communication. Both the Director General of Statistics Estonia and the leading figures of the census gave interviews to all media channels. News was published in daily newspapers, local publications and social media. The leaders of Statistics Estonia and the census met opinion leaders and gave several presentations to politicians. The Estonian Statistical Society dedicated its conference to census-related issues; the topic was also discussed at other scientific conferences and seminars. Several press conferences were held and the Director General, the Project Manager and other members of the census team communicated with various audiences. It was crucial to present the course of enumeration on Statistics Estonia's website in real time. Each hour, the census results were updated for the whole country and the counties (both as absolute figures and ratios).

At the beginning of the census, the media's attitude combined the perception of a significant news item on the one hand, and slight scepticism on the other hand – how many of us are actually left? Can you ensure adequate data protection? As time went by, the attitude became more positive. The website showing census results worked like magic. People became excited about the topic – can we surpass our neighbours in terms of participation in the online census? Will we rank first in Europe? In the world? Interestingly enough, these ambitious goals were not proposed by Statistics Estonia – for example, the goal of achieving a world record was voiced by sociologist Juhan Kivirähk. Yet, enthusiasm was

noticeable everywhere: elderly ladies at the café, young men in tracksuits on the tram, yuppies on a coffee break between meetings – the census, especially the e-census, was always topic number one. The topic peaked on 2 February when the e-census ended and it was apparent that an unofficial world record had been achieved.

Two-thirds of the people were enumerated online, but there were one-third of the people left, including those who were more reluctant and less cooperative by nature. This was when the traditional door-to-door census started. The magic was lost, the media (particularly social media) published stories about minor conflicts, the enumerators' disappointment, the umbrage of people left out of enumeration, but the overall tonality remained positive until the end of the census. The website promoting the e-census failed during the interview census – due to a technical error, it showed remarkably lower population figures for a week, which gave rise to speculations that the number of people living in Estonia has radically declined. These failures were fixed and the errors clarified, the positive attitude persisted. Although the final population number was half a percent less than the expected 1.3 million, it was not a real shock to anyone. The census was generally successful and that message was accepted. In fact, this has been said after every census, but the public has not always shared that opinion.

The census-related publicity did not stop after the census ended. Further activities were targeted at the publication, commenting and clarification of the census results. The results were published by topic for more than a year after the first results had become available, and a press conference was held on the publication of each topic. Such press conferences were frequently supported by radio and TV interviews, newspaper articles and statements in social media. Thus, the PHC media campaign lasted from 2009 to 2013. Several experts have referred to the 2011 Population and Housing Census as one of the most successful national events in terms of maintaining a positive attitude in the society.

INITIAL ASSESSMENT OF THE QUALITY OF PHC 2011

Documents on assessment of quality

Eurostat published the main document on census quality in late 2010:

Commission Regulation (EU) No 1151/2010 of 8 December 2010 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission. *Official Journal of the European Union L 324, 9.12.2010, p. 1–12.*

In addition to the initial guidelines worded on 8 December 2010, Eurostat has been continuously sending further documents to specify and clarify the quality report requirements. The last five guidelines on quality requirements (with an average volume of *ca* 20 pages each) were received on the following dates: 21.05.2013, 13.06.2013, 5.07.2013, 30.07.2013, 5.08.2013, and 28.10.2013, with the last document still being identified as a version, not the final decision.

The quality report also involves the so-called quality hypercubes (35 in total), i.e. tables used for assessing output quality. Considering that this is the first project of this kind (both in terms of hypercubes and quality hypercubes), there is also no unanimity among Eurostat officials as to how to complete the quality hypercubes that measure compliance with the confidentiality requirements of the output hypercubes.

For Estonia, the reviewed quality requirements were established by the PHC Council (Table 6, p. 57). They concerned the census coverage in terms of both persons and dwellings.

Used data sources

PHC 2011 utilised the following data sources, accepted by Eurostat based on Article 4 (1) of the Regulation (EC) No. 763/2008:

- individual enumeration
- other (external) data sources, registers (see p. 7)
- survey data (used only for one characteristic – the number of hours a person normally works in a week – which is not a core characteristic).

During the census preparation period, the registers underwent a thorough analysis and verification of data compliance with the census definitions, of conformity in and between registers, and of updating interval. In most cases, the results met the expectations – persons could be identified well, whereas the identification of addresses and their conformity to the ADS standard left much to be desired.

Quality indicators defined by Eurostat

The quality requirements for statistics published by the EU Member States are briefly summarised as the following quality indicators:

- relevance,
- accuracy,
- timeliness and punctuality,
- availability and clarity,
- comparability,
- coherence.

Obviously, the population and housing census – as a large-scale statistical survey – should also meet all quality indicators.

Relevance

Relevance essentially stands for consumer satisfaction.

In the preparation of the PHC 2011 questionnaire, great attention was devoted to the requests of the customers. It involved holding a series of joint seminars for consumers (representatives of ministries and research institutions). As much as possible, the proposals made during the pilot census (more than 5,000) were also taken into account. At the consumer's request, the questionnaire was supplemented with several characteristics, e.g. grandparents' place of birth

and ability to speak a dialect (asked for the first time), and – under the command of foreign languages – an option to list up to 20 foreign languages.

On the one hand, output planning took into account the former preferences of the consumers, and on the other hand, the opportunities arising from new technologies (hypercubes, detailed tables containing as many characteristics as possible). In view of the needs of the local governments, all significant characteristics are published on the level of the local government units and to some extent, on the level of settlement units. The tables are equipped with detailed explanations (definitions and methodology) in Estonian and in English.

Accuracy

Accuracy characterises the statistical correctness of data with the help of seven indicators.

The variation coefficient (A1) mainly refers to the rate of divergence of the numeric characteristics. It is essentially not applicable for the data set of the population and housing census.

Response rate (A2; V/N, where V is the number of items enumerated/responded and N is the number of items subject to mandatory response). This indicator is of great importance and most efforts related to census quality aim at improving it. In the case of the population and housing census, there are two sets of items subject to enumeration – people (permanent residents of the country) and dwellings (occupied and unoccupied rooms suitable for all-year-round habitation, and occupied non-dwellings).

Since a census is an exhaustive survey, the response rate also measures **coverage**, which is the most significant indicator of census quality today. Coverage is the ratio (K/N) between the persons covered (K) and the number (N) of persons subject to enumeration (target population), usually expressed in percentage. The ratio of persons not covered (N – K)/N represents **under-coverage**. Today, under-coverage poses a serious problem for censuses in all countries.

Under-coverage is caused by:

- the mobility of people, including frequently staying away from home (and abroad), often also the lack of traditional family relationships;
- people's desire of privacy, unwillingness to submit personal data, fear of data leaks;
- the insufficient efforts of some enumerators in finding all persons;
- the special situation in Estonia – the provision of erroneous addresses on the Internet, which were not visited by enumerators, resulting in people being left non-enumerated at those addresses.

Besides the census population K (the number of persons enumerated), the assessment of coverage, under-coverage and over-coverage also requires the target population, i.e. the number (N) of all persons subject to enumeration. There are several options for assessing it, including a follow-up survey, which provides a basis for the statistical evaluation of under-coverage and the target population. An option that is suitable for Estonia is to estimate the total population by using registers.

In recent years, the register system in Estonia has been subject to detailed analyses and constant improvement. The registers belonging to that system were used to a significant extent in the preparation for the census (in the compilation of enumeration lists). The current level of registers allows even more – **they allow determining the population of Estonia with maximum accuracy**. On average, every Estonian resident leaves a mark in three registers per year, but this number increases remarkably when including the sub-registers (e.g. the Health Insurance Fund has several dozen sub-registers according to insurance categories). Therefore, registers provide a rather extensive data set to decide whether a person resides and is actively engaged in Estonia or not. Yet, this data set is not perfect. On the one hand, there are people who are residents of Estonia and were enumerated here, but whose activity was not recorded in any register in 2011. On the other hand, there are also people who have left Estonia but continue to use some of the benefits provided by the Estonian state. Thus, it is necessary to exercise caution when determining the permanent population of Estonia based on registers. It has to be considered that, like any statistical evaluation, such an estimation of the population inevitably contains an error, which needs to be taken into account to provide an accurate estimation.

Who are these people, left out of enumeration but potentially still considered permanent residents of Estonia? Firstly, a person should have an Estonian personal identification code, which allows checking whether the person is entered in any register. It is reasonable to analyse people who – according to the Population Register data – are Estonian permanent residents, but were left out of enumeration during PHC 2011. In principle, there are two possible approaches to decision-making – the so-called soft method, based on expert assessment, or the hard method, based on mathematical statistics.

Expert assessment. After a comprehensive analysis of the registers, the experts establish the combinations of registers that are more likely to indicate where a person permanently resides in Estonia and, based on that, prepare a rule of decision (a so-called index). An index value is calculated for every potential resident. The most common example of an

index is the number of registers containing traces of activity of the person in 2011. Persons with active entries in several independent registers in 2011 are considered permanent residents.

Discriminate analysis is a method of mathematical statistics intended for dividing items into predefined groups. This method suits well for solving this task – i.e. for distinguishing between the residents and non-residents. Another approach similar to a discriminate analysis is distinction by a logistic rule that determines the probability of including each item in one or the other group. The advantage of mathematical statistical methods is that they also allow estimating the accuracy of the relevant estimation.

Assessment of the under-coverage of persons. In a combination of several decision-making strategies, it became evident that the estimated under-coverage is at least 2.1-2.2% for PHC 2011, and residents will include only the persons who get matching results by different methods. In the case of this procedure, the estimated error rate is *ca* 5%, and the possibilities of regarding a non-resident as a resident are limited. More details on the procedure and the results from the assessment of the under-coverage of persons have been published in 2012 in issues No. 3 and 4 of the Quarterly Bulletin of Statistics Estonia.

The assessment of the under-coverage of dwellings principally follows the same method; the estimated under-coverage of dwellings is 1%.

Response rate of a single characteristic (A3; V_i/N_i, where v_i is the number of enumerated/responded items for character i and N_i is the number of items subject to mandatory response) describes single characteristics. It is crucial to reach a high response rate for significant (mandatory) characteristics (approximately 100%). In the case of some characteristics, it was possible to leave the question unanswered (do not wish to answer, do not know). In this census, the only question with the response option "Do not wish to answer" was the question about religion (Do you have any religious affiliation?). However, several questions could be left unanswered by skipping them; still, this was more complicated in the e-census. During the interviews, in the case of most questions, the respondents had three options for refusing to answer: "Do not know" (encoded with the value -2), "Do not wish to respond" (encoded with the value -3) and skipping a question (gap in the database). The response rate of single questions before and after the replacement of missing values (where possible) is shown in Tables 2 and 3 of Annex 1 (pp. 67–68).

On average, the interviews provided information about 96.73% of the questions asked in the Personal Questionnaire. However, this does not include the pre-filled question regarding the place of residence at the time of PHC 2000, in which case the average response rate was 97.8% (lack of response 2.2%). After using register data to fill the gaps, the response rate increased to 98.35% (by 0.55%). Thus, at the end of data processing, the final percentage of missing responses in the database of personal data was 1.65%.

Interviews provided information about 97.92% of the questions in the Dwelling and Household Questionnaire. The response rate increased to 98.68% (by 0.76%) after filling the gaps based on register data, which means that 1.32% of the questions remained unanswered.

Characteristic's imputation rate I_i and ratio (A4a; I_i/N_i, where I_i is the number of post-filled values of characteristic i and N_i is the number of items subject to mandatory response with regard to this character). In PHC 2011, only the foreign sources, i.e. registers, prescribed in the census methodology (combined methodology) were used to fill the gaps, and in some special cases also additional information received during the census (comments in the questionnaire) was used as well. Statistical imputation was not used. Information about the supplemented characteristics is provided in Tables 2 (persons) and 3 (dwellings) of Annex 1(pp. 67–68).

All in all, the average extent to which external data sources were used to supplement the data was 0.55% for personal data and 0.76% for dwelling data.

Item imputation rate I and ratio (A4b; I/N, where I is the number of post-filled items and N is the number of items subject to mandatory response). In PHC 2011, no items were added to the census population either with the help of statistical imputation or external sources. All items in the output tables have been enumerated.

Over-coverage rate U/N and the rate of classification errors (A5), where U is the number of items not subject to enumeration. There was no over-coverage thanks to the identification of persons on the basis of personal identification codes. Theoretically, over-coverage may occur in the case of some persons (less than 1,000) without a personal identification code, but it can still be assured that there is virtually no over-coverage of persons. The over-coverage of dwellings was prevented by organising the addresses according to the address standard (containing also spatial coordinates). This required thorough additional work in the data-processing stage.

Rate of geographical under-coverage (A6). The assessment of under-coverage revealed that it is rather homogenous across the entire country, there were no areas with particularly extensive under-coverage found neither in the e-census nor in the analysis of the distribution of non-enumerated persons (Map 2, p. 35).

Average extent of corrections (A7) is not applicable to the census.

Criteria for evaluating accuracy were specified for Estonia by the PHC Council and they mainly concern the indicator A2 (Table 6, p 57).

Table 6. Quality criteria for the Population and Housing Census 2011[a]
(percentages)

	Very good	Good	Satisfactory	Poor
Coverage rate				
Persons	100	98	95	< 95
Dwellings	100	99	98	< 98
Participation rate of e-census				
Persons	≥ 25	15	10	< 10

[a] Approved at the meeting of the PHC project team on 28.10.2008 and presented at the PHC Council meeting on 26.11.2009.

Timeliness and punctuality

Punctuality is characterised by three indicators:

- punctuality in adherence to the release calendar,
- the period from the end of the survey period to the publication of preliminary results,
- the period from the end of the survey period to the publication of final results.

The general schedule for the publication of results is approved on the international level. The preliminary results were published according to the schedule – two months after the end of the census, on 31.05.2012. The results were published gradually (in 19 parts) according to the release calendar (Table 7). The publication of each section involved a press release, and in most cases also a press conference, introducing and interpreting the content of the tables added to the database. In total, 312 tables (mostly 4–5 characteristics, often even more) were added to the database. The last tables were added to the Statistical Database two years after the end of the census, in March 2014.

Table 7. Schedule for the publication of the PHC 2011 results

31.05.2012	Enumerated permanent residents (preliminary data)
17.08.2012	Population by citizenship (preliminary data)
30.08.2012	Mother tongue and the command of a local dialect among the population with Estonian as the mother tongue (preliminary data)
17.09.2012	Population by ethnic nationality (preliminary data)
12.12.2012	Location, gender and age composition of the population (revised data)
19.12.2012	General data of the population's economic activity
19.03.2013	General data of conventional dwellings
25.03.2013	Economic activity of the population
27.03.2013	Education of the population. Study migration
24.04.2013	Demographic indicators of the population
29.04.2013	Ethnocultural indicators of the population
22.05.2013	Other place of residence. Population temporarily staying in Estonia.
22.05.2013	Place of birth. Persons having changed their country of residence.
22.05.2013	Migration
21.06.2013	Population with native background. Duration of residence and former place of residence. Place of residence at the time of the 2000 Population Census
30.07.2013	Health of the population
28.08.2013	Households and families
25.10.2013	Living conditions of the population and households
26.03.2014	Population in localities

Accessibility and clarity

Accessibility and clarity are characterised by three indicators:

- the number of published and/or sold publications,
- the number of database inquiries,
- the rate of metadata supply.

The amount of printed publications has substantially decreased compared to previous censuses, in the light of giving preference to electronic channels that represent the most suitable means for presenting aggregated data sets. There are two publications focused directly on the census: the bilingual collection "Pilte rahvaloendusest. *Census Snapshots*" (28.06.2013) and "Rahva ja eluruumide loendus 2011. Ülevaade Eesti maakondade rahvastikust" (online publication, 6.05.2013). Articles on the census and census results have been published in the magazine "Eesti Statistika Kvartalikiri. Quarterly Bulletin of Statistic Estonia"*,* and a series of other journals and publications in Estonia ("Akadeemia", "Riigikogu

Teataja", "Eesti Arst", "Papers on Anthropology", booklets of the Estonian Statistical Society, etc.). Additionally, the census has been subject to several blog entries and newspaper articles, including a series of eight articles in the culture newspaper "Sirp".

The number of database inquiries is constantly changing. The usability of the database is also described by the number of tables (312) and the average number of characteristics presented therein. It is also important that besides the directly measured characteristics, these tables contain new characteristics based on the measured ones (origin, command of the Estonian language, de facto marital status, etc.), which significantly facilitates using the tables (see Annex 1 Table 1, p. 64).

All tables are equipped with metadata in Estonian and English, and all concepts have also been defined and made available in the Statistical Database.

Comparability

Comparability is characterised by four indicators:

- the length of compared time series,
- the number of compared time series,
- difference from European standards,
- asymmetries in mutual statistics.

The time series of census data are not long; in most cases they only contain two points in time (2000 and 2011). The main reason for that is that previous censuses (a total of 10 in Estonia) were all based on different regulations, methods and ideology. As the results of the previous censuses have been published in print, it is also possible to use longer time series (6-11 points) for analysing significant characteristics (population, gender and age distribution, ethnic nationality, area of activity, location, urbanisation, etc.).

For the last two censuses, the number of comparable time series reaches 80, taking into account both the characteristics measured in the census and those derived from the measured characteristics. There are no comparable time series available in a physical form, but it is easy to compile them based on the parallel use of the electronically submitted data of the two censuses. Commenting on such time series (data pairs illustrating the dynamics of the results) represented a significant part of the information disclosed at the press conferences dedicated to the publication of census results.

There is very little difference from European standards in the data. In order to ensure comparability, almost all of the core characteristics of PHC 2011 follow Eurostat rules and standards. The only exception is the relations in the household and family. Estonia does not feature all household relations present in Europe – it is not possible to register non-marital cohabitation here and same-sex marriage is not legal either. Therefore, the list of possible household and family types is shorter.

In a few cases, the data of PHC 2011 have been compared to the data of other surveys, but there is still a lot of work to be done in this area. The current problem is that in the case of survey data, the extension factors are not calculated based on the census population (which is under-covered, as we know). A more thorough analysis is provided for employment, which showed poor coincidence in the case of some characteristics. The average estimation of the disability-free life expectancy, on the other hand, showed a rather good coincidence (1-year difference) on the country level with the results gained from sample surveys. The difference is partially due to the fact that institutional households are left out of sample surveys; there were also slight differences in the wording of questions.

Coherence

Coherence is characterised by the rate of data that allow secondary use.

The population and housing census data allow extensive use for very different purposes. They have been introduced to the members of the Riigikogu, the President, representatives of ministries and institutions of higher education, various non-governmental organisations and stakeholders. Population-related questions are still in the centre of public interest, stimulated by the use of census data as argumentation in various discussions.

QUALITY ASSURANCE FOR PHC 2011

Steps taken during census preparation

A series of measures to assure the quality of PHC 2011 were developed based on the experience gained from the pilot census and previous analyses.

- The questionnaire was monitored and tested for user-friendliness by using the best solutions. All questionnaires were available in three languages (Estonian, Russian, English), or even four languages – paper questionnaires had a Finnish version as well. It was possible to use sign language interpreters.

- The questionnaire underwent a cognitive analysis in order to determine the comprehensibility of the questions; the results of the analysis were used to improve comprehensibility.

- The questionnaires were provided with detailed help texts (enabling a more detailed examination of the question background for the e-census participants).

- The questionnaires were equipped with various logic checks to prevent erroneous responses. Soft checks (for unlikely responses) led the respondents to check the accuracy of their response, and strict checks drew attention to obviously erroneous responses that had to be corrected.

- Soft checks did not allow skipping questions easily and drew the respondent's attention to any questions that were left unanswered.

- The enumerators received detailed instructions and passed a five-day training course. The training consisted in an imitation of the actual census situations and the course ended with an examination.

- A support network was developed to assist both self-enumerators and enumerators in terms of both the methodology (related to the content of the questionnaire) and technical issues (login, navigation in the questionnaire, etc.).

- Census tracking (monitoring) was also used, which allowed the census organisers to discover weaknesses and lags in the work of the enumerators, and to support them by means of additional resources as necessary.

- The enumerators' equipment was supplemented with GPS devices for a more detailed determination of the spatial coordinates, helping to find the exact location of the residential buildings enumerated during the e-census, and to pinpoint the location of unoccupied dwellings.

- Various additional sources were used for enumerating unoccupied dwellings and the residents of institutional households.

- The purpose of notification was to
  - encourage people to participate in the census, including the e-census;
  - in the case of the e-census, ensure a relatively equal distribution of the enumeration items over the census period; hence, a counter was set up on the census website, indicating the load on the census environment;
  - recommend people to answer the enumerators or fill their questionnaires on their own;
  - call on local governments to organise address data.

Steps taken during the census

- In terms of the e-census, the most important thing was to activate the people. This was supported by the enumeration information available on the website.

- An absolutely crucial step was to increase the technological capacity of the communication channels after the first failure on 2 January (login problems within half an hour due to an overload when completing the questionnaires took twice as much time as usual). The measures taken were effective and no failures occurred from then on.

- During the census, the security of all channels was controlled to prevent any hacking instances. This information was also passed to media representatives.

- In the course of the census, the technical team removed all software errors (mostly minor corrections that did not have a significant impact on the census results, but made responding more user-friendly).

- Upon the occurrence of problems, the providers of support services were informed in order to find potential solutions to the issues.

- Due to high enumeration activity in the last days of the e-census, the census period was extended by one day (26 hours).

- The enumerators were informed about any problems that occurred during the e-census.

- For better coverage, the second stage of enumeration was organised during the last few days of the interview census (within the last 10–15 days), when the enumerators paid a second visit to all problematic dwellings (e.g. dwellings with the suspicion of a non-enumerated person or entire household living there). The necessary resources were provided by the successful e-census.

Steps taken after the census to improve data quality

After a census, the data gained from the questionnaires has to be organised. On the one hand, this means the adequate use of census information (the main principle for all censuses is reliance upon the information submitted by the enumerated persons), and on the other hand, it involves finding discrepancies in the data set and eliminating them in the best possible way. Thanks to logic checks used already during the (on-line) census, there were not many questionnaires that needed to be corrected. However, a number of inaccuracies resulted from the fact that the respondents had to indicate their address on their own for the first time (in the e-census).

Dependence of the quality of the personal questionnaire on the method of enumeration

Hypotheses

Considering that the application of the e-census to such a great extent is unique in the global enumeration practice, the analysis of the quality of census results by enumeration method is of great interest. There are two possible hypotheses:

- The questionnaires are completed poorly during the e-census due to the lack of verification by the enumerators; also, the person subject to enumeration may not understand the question or even resorts to tomfoolery and writes down answers that make no sense;
- During the e-census, the questionnaires are completed better, because the persons filling them in are more educated and smarter.

Evaluations depending on the method of enumeration can mostly be made on the basis of primary data (the so-called raw data), because after the duplicate analysis, the questionnaires filled by using different methods are combined. However, seeing that cases where electronically completed questionnaires were duplicated with an interview were rather infrequent, the conclusions apply to the final results as well.

The coverage of single characteristics was assessed based on raw census data. At the end of the census, the database of raw personal data contained 1,487,526 entries, but after the first phase of processing the duplicates, it dropped to 1,389,873 personal questionnaires. These provided the basis for the analysis below.

Table 8. Distribution of persons by enumeration method

Person's status in household	Enumeration method				Total
	laptop	paper	online	combined	
N/A	83	3	80	1	167
Permanent residents	437,919	2,072	859,845	6,106	1,305,942
Temporary residents	15,799	40	28,600	555	44,994
Residents who have left	19,223	28	19,297	222	38,770
Total	473,024	2,143	907,822	6,884	1,389,873

The comparative analysis concerns only the questionnaires of permanent residents, totalling 1,305,942 (93.96% of all questionnaires). Only permanent residents constitute the population. For the rest of the enumerated persons – temporary residents (3.24%) and those who had left Estonia (2.79%) – short questionnaires were filled out. Henceforth, the online questionnaires include those filled by the combined method, where the interviewer supplemented the questionnaire that was first filled out online. The share of such questionnaires among all questionnaires is less than 0.5%. Paper questionnaires – only 0.15% of all filled questionnaires – were naturally filled out in the course of the interview census (in some cases, by phone interview) and shall not be considered separately.

Duplicate questionnaires have not been removed from this data set. The processing of duplicates would probably improve the coverage of single characteristics, because the so-called original questionnaire which is stored in the database retains the information included in all duplicates.

Sex and age. Sex was not indicated in 13 questionnaires, age in 16 questionnaires. This reduces the number of questionnaires analysed in Table 9. One questionnaire indicated the age of 111 years, which apparently could not be

accurate; the rest of the responses caused no suspicion. Thus, the coverage of basic characteristics was very good. The missing information can be supplemented with the help of personal identification codes.

The coverage of single questions by different enumeration methods. This section considers personal data without identification data (given name and surname, personal identification code) and a substantial part of address data. Both of these data categories are treated separately.

Table 9 shows the percentage of missing answers per each single characteristic (question) by response method, taking into account the persons who had to answer that question. "Refusal" includes two response options: "Do not wish to reply" and "Do not know". In general, that option was unavailable for those who responded online.

Table 9. Comparison of response methods by coverage of single characteristics

Questionnaire item	Laptop		Paper		Internet		Combined		Total	
	refusal	no response	refusal	no response	refusal	no response	refusal	no response	refusal	no response
Type of dwelling	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.00
Respondent	0.00	0.05	0.00	0.72	0.00	0.07	0.00	0.05	0.00	0.06
Secondary place of residence	0.15	0.25	0.28	1.79	0.00	0.23	0.36	0.34	0.10	0.24
Legal marital status	0.31	0.06	0.01	0.73	0.00	0.11	0.33	0.12	0.11	0.09
Has given birth	0.25	0.08	0.01	1.01	0.00	0.11	0.43	0.14	0.09	0.10
Number of live-born children	0.06	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.02	0.02
Country of birth	0.32	0.06	0.28	0.97	0.00	0.12	0.51	0.15	0.15	0.10
Mother tongue	0.14	0.05	0.26	0.63	0.00	0.12	0.29	0.11	0.09	0.10
Ability to speak a dialect	0.21	0.02	0.03	0.63	0.00	0.02	0.18	0.05	0.07	0.02
Dialect spoken	0.01	0.00	0.00	0.00	0.00	1.17	0.00	0.00	0.00	0.69
Command of foreign languages*	0.31	0.07	27.80	1.01	0.00	0.13	0.59	0.18	0.16	0.11
Foreign languages spoken	0	0.00	0	0.00	0	0.03	0	0.02	0	0.02
Existence of religious affiliation	0.48	0.07	0.26	1.04	0.00	0.30	0.51	0.21	0.22	0.22
Indication of religious affiliation	0.10	0.10	0.00	0.31	0.00	0.02	0.00	0.00	0.05	0.06

As the characteristics concerning education and employment need converting and re-coding, their quality cannot be assessed based on primary data alone and they are thus not included in the table.

The online census did not allow direct refusal to respond, although it was possible to skip a question. It appears that with the majority of comparable questions, the percentage of those who did not respond or refused to respond was lower in the case of the online census. The only exceptional characteristic is the spoken dialect – the rate of non-response was more than one percent in the case of persons who were asked about it in online questionnaires. The number of those using other enumeration methods was so low that it did not have a significant impact on the final result. However, it still shows that the combined response method did not give better results than the online census or the interview census separately and, compared to the above-mentioned methods, the paper questionnaire yielded the most modest result – as was excepted since it was partially a phone interview, which was carried out according to a somewhat shorter programme during the last census days.

Indication of addresses. The biggest problems in ensuring data quality were address-related. It was for the first time in the history of Estonian censuses that respondents had to indicate their address on their own. The ADS address system had been developed right before the census, it had not quite been established yet and people were not used to that standard. Thus, organising the address data was rather time-consuming. Although the standard prescribes using the names of small places (cottage and gardening association) on ADS level four, some of them ended up in the street name field; street names (ADS level five) and numbers (ADS level seven) are combined on ADS level six (name of farm or land unit). Both the online respondents and interviewers had problems with entering street names containing a personal name. In the case of such street names, the ADS requires indicating the given name of the person (such as Albert-August Tiimann Street in Narva, Eduard Vilde Road in Tallinn, etc.), but people are generally not used to it. In some cases they do not even know that the street name is a personal name.

Problems were also encountered in the case of new and multi-sectional buildings. As the Survey Fieldwork Information System (VVIS) required an exact match for each building and address, numbers (1), (2), etc. were attributed to the house number in the list of buildings to denote the separate parts of the house.

It was somewhat surprising that there were more address errors to be fixed in the questionnaires of the interview census than in those of the e-census. In the course of the census, 127,267 address operator cases had to be solved, among them, 59,231 from the web and 68,036 from the interviewers.

Another common problem is that many people (up to 20%, according to estimations) have not registered their actual place of residence, and even during the online census some of them indicated their registered place of residence instead of their actual place of residence, which was one of the reasons for under-coverage.

Indication of personal identification codes. For various reasons, approximately 10% of the e-census respondents and approximately 40% of the interview census respondents did not indicate their personal identification codes. Identifying them meant additional work for the operators, and it also obstructed the fast publication of census data. Personal identification codes were necessary for sorting out the duplicates. It occurred that for various reasons, 8% of the enumerated persons had been enumerated several times (2.1 times on average), and most frequently it happened to persons temporarily residing away from their regular place of residence. After the primary data processing was over, 1,384 persons or 0.1% of those enumerated remained without a personal identification code.

Overall data quality

In the course of data collection, both the e-census and the interview census utilised various logic checks that prevented illogical (non-compliant) responses, and therefore the overall quality of the census data is good or very good. For the majority of the questions, non-response was also excluded (or made difficult), resulting in very good coverage of most characteristics. According to the encoders, there was a relatively low level of textual nonsense (tomfoolery). A comparison of the data received from the e-census and the interview census revealed that:

- in the e-census, all texts (names, names of employers/ institutions, activities, job titles, etc.) were written more accurately and correctly;
- during the interview census, there was less intentional tomfoolery.

As for the e-census, considering the better coverage of single characteristics, the remarkably higher proportion of indicated personal identification codes and fewer errors in address data, it can be asserted that the responses received from the e-census generally have better quality than those received from the interview census. The main reason for that is most likely the different contingent of respondents.

What to do with persons who were not enumerated?

As the publisher of census results, Statistics Estonia has three options in the case of under-coverage:

- ignore under-coverage and consider the census result as the actual population – this has been done in most countries and most censuses until the present day;
- verify coverage and assess under-coverage, but still consider the census result as the actual population – this was done in Estonia in 2000 and has been done in other countries in the last census rounds;
- verify coverage and adjust the census result by including the non-enumerated persons in the actual population in addition to those enumerated. This is accepted on the international level, but has not been used much due to lack of relevant international methodology.

The first option was inevitable in earlier periods in history, because under-coverage and over-coverage balanced each other out. Today, however, under-coverage represents the main problem with census quality.

The verification of coverage and the correction of coverage errors, i.e. adding non-enumerated persons to the estimated population, are possible if the country in question has a well-functioning system of compatible (linked) registers, the data of which are updated and specified in a timely manner.

Just like the previous time, the second option was chosen for PHC 2011. Although the public was informed of under-coverage, all the census data are published based on the census population, i.e. by taking into account only the persons who were actually enumerated. Subsequent population statistics, however, started using the adjusted population number. Additionally, the data from PHC 2000 and the population number of 2000–2011 were revised as well. According to estimations, the population number of PHC 2000 was also under-covered by 2.2%.

SOURCES

- Article 2(8) of Commission Regulation (EU) No 1151/2010 (2010).
- Census methodology: Key results the UNECE Survey on National Census Practices, and first proposals about the CES Recommendations for the 2020 census round. Note by the UNECE Task Force on Census methodology. (2013). UNECE. [www] http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2013/census_meeting/3_E_x_15_Aug_WEB_revised_map.pdf
- Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing. (2008). Official Journal of the European Union, L 218, 13.8.2008, pp. 14–20.
- Statistical Database. [online database] http://pub.stat.ee/px-web.2001/I_Databas/Population_Census/databasetree.asp (20.12.2013)
- Tiit, E.-M., Meres, K., Vähi, M. (2012). Assessment of the target population of the census. – Eesti Statistika Kvartalikiri. Quarterly Bulletin of Statistics Estonia, No. 3, pp. 79–108. Tallinn: Statistics Estonia.
- Tiit, E.-M. (2012). Assessment of under-coverage in the 2011 Population and Housing Census. – Eesti Statistika Kvartalikiri. Quarterly Bulletin of Statistics Estonia, No. 4, pp. 110–115. Tallinn: Statistics Estonia.
- Tiit, E-M., Vähi, M. (2012). Enumerators' activity after the Census. – Eesti Statistika Kvartalikiri. Quarterly Bulletin of Statistics Estonia, No. 2, pp. 102–111. Tallinn: Statistics Estonia.
- Tiit, E-M. (23.08.2012). Loendusmeeskond jätkab tööd. – Sirp.
- UNECE 2006. Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing. United Nations, New York and Geneva.

Annex 1. Tables

Table 1. List of PHC 2011 characteristics with sources

	Question/characteristic	Number in question-naire	European code	Purpose	Unit	HC	Source
1	Given name	A01		Identifying	Person		Census
2	Surname	A02		Identifying	Person		Census
3	Personal identification code[a]	A03		Identifying	Person		Census
4	Sex	A04	SEX	Output	Person	+	Census
5	Year of birth	A05A		Output	Person		Census
6	Month of birth	A05B		Auxiliary	Person		Census
7	Day of birth	A05C		Auxiliary	Person		Census
8	Age	A05D	AGE	Output	Person	+	Census
9	Country of permanent residence			Determining	Person		Census
10	City/county of residence	E01	GEO	Output	Person	+	Census
11	Settlement of residence	E01		Output			Census
12	Address of residence	E01		Identifying			Census
13	Size of locality		LOC	Output	Person	+	Derived E01
14	Status of person (permanent resident/non-permanent resident)	AX1		Identifying	Person		Census
15	Respondent (person him-/herself or somebody else)	A00		Auxiliary			Census
16	Existence of secondary residence	A07B		Output	Person		Census
17	Length of stay at secondary residence	A09A		Output	Person		Census
18	Country of secondary residence	A08B		Output	Person		Census
19	City/county of secondary residence	A08A		Output	Person		Census
20	Legal marital status	A10	LMS	Output	Person	+	Census
21	De facto marital status			Output	Person		S_matrix
22	Legal/de facto marital status			Output	Person		S_matrix, A10
23	Status in family		FST	Output	Person	+	S_ matrix
24	Status in household		HST	Output	Person	+	S_ matrix
25	Type of family		TFN	Output	Family	+	S_ matrix
26	Type of private household		TPH	Output	Household	+	S_ matrix
27	Size of family		SFN	Output	Family	+	S_ matrix
28	Size of private household		SPH	Output	Household	+	S_ matrix
29	Has given birth	A11		Output	Person		Census
30	Number of children given birth to	A11A		Output	Person		Census
31	Age at the time of birth of the first child			Output	Person		SE register
32	Country of birth	A13	POB	Output	Person	+	Census
33	Place of birth	A13A		Output	Person		Census
34	Father's country of birth	A14B		Output	Person		Census
35	Mother's country of birth	A14A		Output	Person		Census
36	Parents' country of birth			Output	Person		A14A, A14B
37	Grandparents' country of birth	A14C		Output	Person		Census
38	Native origin			Output			A13,A14A, A14B,A14C
39	Citizenship	A15	COC	Output	Person	+	Census
40	Ethnic nationality	A17		Output	Person		Census
41	Mother tongue	A19		Output	Person		Census
42	Ability to speak a dialect	A19A		Output	Person		Census
43	Dialect spoken	A19B		Output	Person		Census
44	Command of foreign languages	A20		Output	Person		Census
45	Foreign languages spoken	A20A		Output	Person		Census
46	Command of Estonian			Output	Person		A20, A20A
47	Religious affiliation	A21		Output	Person		Census

[a] The data of the birth date, age, sex and personal identification code was used to verify data (upon entry); generally, not all of them were asked from the respondent directly.

Table 1. List of PHC 2011 characteristics with sources

Cont.

	Question/characteristic	Number in question-naire	European code	Purpose	Unit	HC	Source
48	Religion	A21A		Output	Person		Census
49	Change in residence	A22		Output	Person		Census
50	Time of last migration	A22A		Output	Person		Census
51	Previous country of residence	A23		Output	Person		Census
52	Previous place of residence in Estonia	A23A		Output	Person		Census
53	Previous settlement of residence in Estonia	A23A		Output	Person		Census
54	Place of residence a year ago		ROY		Person	+	A22, A22A, A23, A23A
55	Experience of living in a foreign country	A24		Output	Person		Census
56	Time of immigration/return	A25A	YAE	Output	Person	+	Census
57	Recent country of residence in a foreign country	A26		Output	Person		Census
58	Country of usual residence in 2000	A27		Output	Person		Census
60	County of usual residence in 2000	A27A		Output	Person		Census
61	Level of general education	A28		Output	Person		Census
62	Literacy	A28A			Person		Census
63	Level of vocational education	A29			Person		Census
64	Education required for entry	A30			Person		Census
65	Secondary specialised education	A31A			Person		Census
66	Vocational secondary education	A31B			Person		Census
67	Vocational and basic education	A31C			Person		Census
68	Higher education	A33			Person		Census
69	Level of education			Output			A28–A33
70	Level of education based on ISCED		EDU	Output	Person	+	A28–A33
71	Level of education being acquired			Output	Person		EHIS
72	Location of educational institution			Output	Person		EHIS
73	Main source of subsistence	A34		Output	Person		Census
74	Working on reference week	A36		Auxiliary	Person		Census
75	Existence of place of employment on reference week	A37		Auxiliary	Person		Census
76	Seeking work	A38		Auxiliary	Person		Census
77	Readiness to commence work	A39		Auxiliary	Person		Census
78	Employment status			Output	Person		A36–A39
79	Earlier employment	A40		Output	Person		Census
80	Year of last employment	A40A		Output	Person		Census
81	Social status of non-working person	A41		Output	Person		Census
82	Socio-economic status		CAS	Output	Person	+	A36–A39, A41
83	Name of place of work	A42		Auxiliary	Person		Census
84	Name of subordinate unit	A42B		Auxiliary	Person		Census
85	Economic activity of company	A43A	IND	Output	Person	+	Census
86	Economic activity of subordinate unit	A43B	IND[i]	Output	Person	+	Census
87	Occupation (ISCO 08)	A44A	OCC	Output	Person	+	Census
88	Description of duties	A44B		Auxiliary	Person		Census
89	Number of subordinates	A45		Auxiliary	Person		Census
90	Labour status	A46	SIE	Output	Person	+	Census
91	Location of job	A47	LPW	Output	Person	+	Census
92	Labour and socio-economic status			Output	Person		A36–A39, 93A41, A46
93	Country of place of work	A47B		Output	Person		Census
94	County, rural municipality, settlement of place of work	A47A		Output	Person		Census
95	Usual number of working hours per week			Output	Person		Labour Force Survey
96	Existence of health problem	A50		Output	Person		Census

Table 1. List of PHC 2011 characteristics with sources

Cont.

	Question/characteristic	Number in question-naire	European code	Purpose	Unit	HC	Source
97	Existence of health restriction	A51		Output	Person		Census
98	Life expectancy			Output	Person		Calculated
99	Disability-free life expectancy			Output	Person		Calculated A51
100	Type of dwelling	E02	TLQ, HAR[ii]	Output	Dwelling	+	Census
101	Occupancy of dwelling	E03	OCS	Output	Dwelling	+	Census
102	Type of building/dwelling		TOB	Output	Dwelling	+	Derived_E02
103	Number of dwellings in residential building			Output			Derived_address
104	Number of households in dwelling	E04		Output	Dwelling		Census
105	Number of residents in dwelling		NOC	Output	Dwelling	+	S_matrix
106	Respondent (housing and household questionnaire)	EX2		Auxiliary			Census
107	Age and number of children			Output	Dwelling		Number of residents, ages
108	Tenure status of dwelling	EL14	TSH	Output	Dwelling	+	Census
109	Owner of dwelling	E15	OWS	Output	Dwelling	+	Census
110	Construction time	E16	POC	Output	Dwelling	+	Census
111	Number of rooms	E17	NOR	Output	Dwelling	+	Census
112	Number of rooms per resident		DRM	Output	Dwelling		Calculated_E17, number of residents
113	Area of dwelling	E18	UFS	Output	Dwelling	+	Census
114	Average area per resident		DFS	Output	Dwelling		Calculated E18, number of residents
115	Existence of kitchen/kitchenette	E19		Output	Dwelling		Census
116	Water supply	E21	WSS	Output	Dwelling	+	Census
117	Bath	E23	BAT	Output	Dwelling	+	Census
118	Sauna	E24		Output	Dwelling		Census
119	Washing facilities			Output	Dwelling		E23,E24
120	Toilet facilities	E25	TOI	Output	Dwelling	+	Census
121	Main heating option	E26	TOH	Output	Dwelling	+	Census
122	Comfort characteristics			Output	Dwelling		E19, E21, E23, E24, E25, E26
123	Existence of other dwellings	EL27		Output	Household		Census
124	Number of other dwellings	EL27A		Output	Household		Census
125	Agricultural production for own consumption	EL28		Output of Agricultural Census	Household		Census

Comments on Table 1

- ▪ The characteristics listed under "Question/characteristic" are from output tables, but also from the questionnaire in some cases.
- ▪ The columns "Number in questionnaire" and "European code" list, respectively, the number of the question in the questionnaire and the agreed three-letter code of core characteristics.
- ▪ The column "Purpose" shows whether a characteristic has been published in Estonian output tables. In some cases, the purpose of a characteristic is to identify an object or determine whether an object is an enumeration object. Some are auxiliary characteristics, which are not published.
- ▪ The column "Unit" marks the object of study, in the case of the census, there are five options here: person, household, family, dwelling, residential building.
- ▪ The column "Source" shows whether a characteristic was taken from the census, calculated/derived based on the characteristics received in the census (in which case, the number of the characteristic used is shown), or taken from a survey (the name of the survey is shown) or from a register (the name of the register is shown).

Table 2. Response rate of personal questionnaires before and after supplementation with register data

Question number	Characteristic	Responses	Non-responses in census	%	Supple-mented	%	Register used	Deleted	Remained unknown	%
A10	Legal marital status	1 094 564	20 847	1,90	7 205	0,66	RR	0	13 642	1,25
A13	Country of birth	1 300 013	27 353	2,10	26 897	2,07	RR, PHC 2000	0	456	0,04
A15	Citizenship	1 300 013	28 291	2,18	27 823	2,14	RR	0	468	0,04
A24	Has lived abroad	1 294 455	21 688	1,68	0	0,00		0	21 688	1,68
A25B	Year of arrival from abroad	236 431	4 825	2,04	0	0,00		0	4 825	2,04
A44A	Occupation (ISCO 08)	617 348	10 274	1,66	0	0,00		0	10 274	1,66
A47	Location of job	557 975	648	0,12	0	0,00		0	648	0,12
A28	General education school	1 094 564	27 432	2,51	11 635	1,06	PHC 2000, EHIS	0	15 797	1,44
A29	Vocational education	1 091 049	28 350	2,60	7 991	0,73	PHC 2000, EHIS	0	20 359	1,87
A30	Education required for entry	419 558	6 982	1,66	6 473	1,54	PHC 2000, EHIS	0	509	0,12
A31A–C	Vocational general education	419 049	22 844	5,45	15 897	3,79	PHC 2000, EHIS	0	6 947	1,66
A33	Higher education	242 792	6 778	2,79	6 778	2,79	PHC 2000, EHIS	0	0	0,00
A36	Working	1 094 564	16 647	1,52	0	0,00		0	16 647	1,52
A37	Absent from work	585 104	16 689	2,85	0	0,00		0	16 689	2,85
A38	Seeking work	536 589	17 685	3,30	0	0,00		0	17 685	3,30
A39	Would commence work	78 599	100	0,13	0	0,00		0	100	0,13
A40	Was working	68 963	234	0,34	42	0,06	PHC 2000	0	192	0,28
A40A	Last employment	59 373	698	1,18	0	0,00		0	698	1,18
A41	Non-working	467 626	18 024	3,85	840	0,18	Ministry of Defence	0	17 184	3,67
A43A	Economic activity of company, EMTAK[a]	617 348	2 982	0,48	0	0,00		0	2 982	0,48
A43B	Economic activity of subordinate unit, EMTAK	212 490	1 395	0,66	0	0,00		0	1 395	0,66
A46	Labour status	617 348	784	0,13	0	0,00		0	784	0,13
A47A	Location of place of work based on EHAK[b]	53 247	49 558	9,31	0	0		0	49 558	9,31
A47B	Country of place of work	557 975	1 342	0,24	0	0		0	1 342	0,24
A07B	Secondary residence	1 294 455	21 706	1,68	0	0,00		0	21 706	1,68
A09A	Length of stay at secondary residence	154 660	2 945	1,90	0	0,00		0	2 945	1,90
A08B	Country of secondary residence	79 461	165	0,21	0	0,00		0	165	0,21
A08A	Location of secondary residence based on EHAK	52 581	83	0,16	0	0,00		0	83	0,16
A11	Has given birth	596 926	5 295	0,89	0	0,00		0	5 295	0,89
A11A	Number of children given birth to	454 855	6 256	1,38	3 652	0,80		0	2 604	0,57
	Age at the time of birth of the first child	454 855	454 855	100,00	439 442	96,61	SE, PHC 2000	0	15 413	3,39
A13A	Place of birth based on EHAK	1 096 857	22 584	2,06	3 205	0,29	RR	0	19 379	1,77
A14A	Mother's country of birth	1 294 455	27 334	2,11	19 842	1,53	RR, PHC 2000	0	7 492	0,58
A14B	Father's country of birth	1 294 455	34 275	2,65	17 534	1,35	RR, PHC 2000	0	16 741	1,29
A14	Grandparents' country of birth	1 294 455	23 052	1,78	11 097	0,86		0	1 195	0,92
A17	Ethnic nationality	1 294 455	14 930	1,15	13 295	1,03	RR, PHC 2000	0	1 635	0,13
A19	Mother tongue	1 294 455	23 934	1,85	22 211	1,72	RR, PHC 2000, EHIS	0	1 723	0,13
A19A	Ability to speak a dialect	853 331	15 001	1,76	0	0,00		0	15 001	1,76
A19B	Dialect spoken	131 239	1 109	0,85	0	0,00		0	1 109	0,85
A20	Command of foreign languages	1 249 434	24 182	1,94	0	0,00		0	24 182	1,94

[a] Estonian Classification of Economic Activities
[b] Classification of Estonian Administrative Units and Settlements

Table 2. Response rate of personal questionnaires before and after supplementation with register data

Cont.

Question number	Characteristic	Responses	Non-responses in census	%	Supple-mented	%	Register used	Deleted	Remained unknown	%
A20A	Foreign languages spoken	856 225	0	0,00	0	0,00		0	0	0,00
A21	Religious affiliation	1 094 564	23 888	2,18	0	0,00		0	23 888	2,18
A21A	Religion	320 872	1 274	0,40	0	0		0	1 274	0,40
A22	Lived there since birth	1 294 455	14 975	1,16	0	0,00		0	14 975	1,16
A22AB	Time of settling in place of residence	828 532	80 964	9,77	0	0,00		0	80 964	9,77
A23	Previous country of residence	828 532	10 056	1,21	0	0,00		0	10 056	1,21
A23A	Previous place of residence based on EHAK	647 931	84 779	13,08	0	0,00		0	84 779	13,08
A26	Recent country of residence in a foreign country	236 431	3 904	1,65	0	0,00		0	3 904	1,65
A27	Country of usual residence in 2000	1 132 176	22 550	1,99	16936	1,50	PHC 2000	0	5 614	0,50
A27A	Place of usual residence based on EHAK in 2000	1 115 254	464 736	41,67	463 756	41,58	PHC 2000	0	980	0,09
A28A	Literacy	3515	85	2,42	0	0,00		0	85	2,42
A34	Main source of subsistence	1 294 455	16 786	1,30	0	0,00		0	16 786	1,30
A45	Subordinates	159 986	914	0,57	0	0,00		0	914	0,57
A50	Health	1 294 455	23 643	1,83	0	0,00		0	23 643	1,83
A51	Health restrictions	1 294 455	24 112	1,86	0	0,00		0	24 112	1,86
	Relationship in household	1 190 712	14 954	1,26	11 808	0,99	RR, PHC 2000	4	3 146	0,26

Table 3. Response rate of dwelling and household questionnaires before and after supplementation with register data

Question number	Characteristic	Responses	Non-responses in census	%	Supple-mented	%	Register used	Deleted	Remained unknown	%
E02	Type of dwelling	662 477	3 355	0,51	0	0,00		0	3 355	0,51
E03	Occupancy of dwelling	662 477	11 337	1,71	0	0,00		0	11 337	1,71
E15	Owner	654 814	8 864	1,35	0	0,00		0	8 864	1,35
E16	Construction time	654 814	61 783	9,44	50 082	7,65	EHR	0	11 701	1,79
E17	Number of rooms	654 814	11 638	1,78	5 033	0,77	EHR	0	6 605	1,01
E18	Area	654 814	9 077	1,39	2 342	0,36	EHR	0	6 735	1,03
E19	Kitchen, kitchenette	654 814	11 044	1,69	4 592	0,70	EHR	0	6 452	0,99
E21	Water supply	654 814	7 217	1,10	879	0,13	EHR	0	6 338	0,97
E23	Washing facilities	654 814	8 494	1,30	356	0,05	EHR	0	8 138	1,24
E24	Sauna	113 094	7 662	6,77	213	0,19	EHR	0	7 449	6,59
E25	Toilet	654 814	8 119	1,24	845	0,13	EHR	0	7 274	1,11
E26	Heating	654 814	8 343	1,27	385	0,06	EHR	0	7 958	1,22
EL14	Tenure status	591 286	9 735	1,65	0	0,00			9 735	1,65
EL28	Agricultural production for own consumption	599 832	10 338	1,72	0	0,00			10 338	1,72

The census data were supplemented with data from six registers:

- The Population Register (RR), 60,305 entries were used for data supplementation;
- The database of PHC 2000, 102,068 entries were used for data supplementation;
- The Estonian Education Information System (EHIS), 39,911 entries were used for data supplementation;
- The database of the Ministry of Defence, 840 entries were used for data supplementation;
- Statistics Estonia's database of births, 76,240 entries were used for data supplementation;
- The State Register of Construction Works (EHR), 64 727 entries were used for data supplementation (construction time, which several respondents did not know, was specified the most).

[i] If a company has subordinate units with a different address or area of economic activity, the activity and location of the subordinate unit are recorded.
[ii] Includes the Eurostat sub-characteristics HAR_DW (living in ordinary housing), HAR_CLQ (living in institutional housing), HAR_H_OTH (living in non-residential dwellings) and HAR_HMLS (homeless).

2011 POPULATION AND HOUSING CENSUS

PERSONAL QUESTIONNAIRE
(Paper version)

ENUMERATOR'S PART

Kpv	DATE OF INTERVIEW	Kpvk DATE OF CONTINUING INTERRUPTED INTERVIEW
A	DAY |__|__|	DAY |__|__|
B	MONTH |__|__|	MONTH |__|__|
C	YEAR |__|__|__|	YEAR |__|__|__|
Alg	INTERVIEW STARTED	Algk TIME OF CONTINUING INTERRUPTED INTERVIEW
A	HOUR |__|__|	HOUR |__|__|
B	MINUTES |__|__|	MINUTES |__|__|
Kviis	TYPE OF SURVEY 1 – LAPTOP SURVEY 2 – PAPER QUESTIONNAIRE SURVEY	
Lik	ENUMERATOR'S PERSONAL IDENTIFICATION CODE I__I__I__I__I__I__I__I__I__I__I__I	
Rg	REGIONAL MANAGER NUMBER |__|	
Rng	DISTRICT HEAD NUMBER |__|__|	
Prk	AREA SUPERVISOR NUMBER |__|__|__|	
Jsk	ENUMERATION AREA NUMBER |__|__|__|__|	
Lnr	SEQUENTIAL NO. OF HOUSEHOLD IN DWELLING I__I__I	
IA	PERSON'S ROW NUMBER IN HOUSING AND HOUSEHOLD QUESTIONNAIRE LIST A |__|__| *Personal questionnaire document no. in FMS I__I__I__I__I__I to be filled in by area supervisor*	
IB	PERSON'S COLUMN NUMBER IN HOUSING AND HOUSEHOLD QUESTIONNAIRE LIST B |__|__| *Personal questionnaire document no. in FMS I__I__I__I__I__I to be filled in by area supervisor*	
Aonr	DATA OBJECT NO. I__I__I__I__I__I__I__I__I__I__I__I__I *to be filled in by enumerator*	

Tallinn 2011

PERSONAL DATA

INSERT HERE PERSONAL DATA A01–A06A FROM HOUSING AND HOUSEHOLD QUESTIONNAIRE AND RE-CHECK THEM.

A01	**Given name**
A02	**Surname**
A03	**Personal identification code** |__|__|__|__|__|__|__|__|__|__|__| MAY BE LEFT EMPTY IF PERSONAL IDENTIFICATION CODE IS ALREADY PROVIDED IN LIST A OR B.
A04	**Sex** 1 Male 2 Female
A05	**Date of birth** |__|__| |__|__| |__|__|__|__| Day Month Year
D	Age as of 31 December 2011 |__|__|__|
A06A	**Place of usual residence** (TO BE FILLED IN FOR A PERMANENT RESIDENT OF THE DWELLING)
AX4	**DIRECTION TO NEXT QUESTION:** For a temporary resident of the dwelling (AX1 = 2) → QUESTIONS FOR TEMPORARY RESIDENTS For a person who has left Estonia (AX1 = 3) → QUESTIONS ON PERSONS WHO HAVE LEFT ESTONIA
A00	**WHO WILL PROVIDE ANSWERS ABOUT THIS PERSON?** **1** PERSON HIMSELF/HERSELF OR HIS/HER CARER/GUARDIAN (IF AT LEAST 15 YEARS OF AGE) **2** PARENT OR GUARDIAN OF THIS CHILD (IF UNDER 15 YEARS OF AGE) **3** MEMBER OF THE HOUSEHOLD **4** OTHER PERSON
AX2	**DIRECTION TO NEXT QUESTION:** If dwelling is missing (E02 = 11 or 12) → **A10**
A07B	**Do you/does he/she stay at another place of residence for some time in a week or year?** *Only consider place of residence located outside the city or rural municipality of your/his/her usual place of residence.* **1** Yes **2** No → **A10**
A09A	**How long is your/his/her annual stay at the secondary place of residence?** ▪ *Include non-consecutive days. For example, if you have/he/she has used the other place of residence on all weekends during the year, the duration of stay at the secondary place of residence is over three months.* ▪ *In case you live/he/she lives at several other addresses beside the place of usual residence, indicate the one with the longest period of stay.* **1** Less than 3 months → **A10** **2** 3–6 months **3** More than 6 months
A08 **MK** **V** **A**	**Where is your/his/her secondary place of residence located?** **In Estonia:** County……………………………………………………………………………………………………….. Local government (city / rural municipality)…………………………………………………………………. C. without municipal status / (small) town / village / c. distr.………………………………………………… or **in a foreign country** (INDICATE COUNTRY)……………………………………………………

Please indicate all the following information as of 31 December 2011.

A10	TO BE ASKED ABOUT PERSONS OF AT LEAST 15 YEARS OF AGE. **What is your/his or her legal marital status?** **1** Never been legally married **2** Legally married **3** Divorced **4** Widowed

QUESTIONS A11 AND A11A TO BE ASKED ABOUT WOMEN OF AT LEAST 15 YEARS OF AGE.

A11	**Have you/has she given birth to children?** *Include all live-born children before 31 December 2011.* **1** Yes **2** No → **A13**
A11A	**How many children have you/has she given birth to?** Number of live-born children: |__|__|
A13 ☼ p 11	**What is your/his/her country of birth?** *Country of birth is your/his/her mother's place of usual residence at the time of your/his/her birth.* OPTION 2 IS NOTED IF THE PERSON WAS BORN ON THE MENTIONED TERRITORY BEFORE 1945. **1** Estonia **2** Former territory of Estonia – Petserimaa (Pechory) or area to the east of the Narva River → **A14** **3** Russia → **A14** **4** Ukraine → **A14** **5** Byelorussia → **A14** **6** Other country (NOTE)………………………………… → **A14**
A13A MK V A Y	**Please indicate your/his/her place of birth in Estonia. Place of birth is your/his/her mother's place of usual residence at the time of your/his/her birth (not location of the hospital).** *If you are unable to answer on the basis of the current administrative division, note the place of birth on the basis of the former administrative division and specify the validity period of that division, as far as you know.* FORMER ADMINISTRATIVE DIVISION County (district) …………………………………………………………………….……….... Rural municipality (village council)………………………………………………………………….. City / town / small town / village.. …………………………………………………………………. |__|__|__|__| Year when the administrative division took effect (in case of former administrative division)
A14 ☼ p 11	**In which country were your/his/her parents born?** OPTION 2 IS INDICATED IF A PARENT WAS BORN ON THE MENTIONED TERRITORY BEFORE 1945. **Mother** **Father** **1** Estonia **2** Former territory of Estonia – Petserimaa (Pechory) or area to the east of the Narva River **3** Russia **4** Ukraine **5** Byelorussia **6** Other country (NOTE) **mother:**………………………….. **father:**…………………………………

3

A14C	**Are your/his/her grandparents born in Estonia? Please state, whether …**
	1 all four grandparents were born in Estonia
	2 some of the grandparents were born in Estonia
	3 none of the grandparents were born in Estonia
	4 the country of birth of all grandparents is unknown?
	NB! Persons born before 1945 on the former territory of the Republic of Estonia (Pechory or area to the east of the Narva River) are also considered to be born in Estonia.

A15	**What is your/his/her country of citizenship?**
	1 Estonia
	2 Russia
	3 Other (NOTE) …………………………………………………………………………….
	4 Undefined citizenship (with an Estonian alien's passport)

A17	**What is your/his/her ethnicity?**		
	1 Estonian	**3** Ukrainian	**5** Finnish
	2 Russian	**4** Byelorussian	**6** Other ethnicity (NOTE)…………………………

A19	**What is your/his/her mother tongue?**		
	1 Estonian	**3** Ukrainian → **A20**	**5** English → **A20**
	2 Russian → **A20**	**4** Finnish → **A20**	**6** Other language (NOTE) ...………..………… → **A20**

QUESTIONS A19A AND A19B TO BE ASKED ABOUT PERSONS OF AT LEAST THREE YEARS OF AGE, WHOSE MOTHER TONGUE IS ESTONIAN.

A19A	**Do you/ does he/she speak some local language form, dialect or subdialect?**		
	1 Yes	**2** No → **A20**	
A19B	**Please name this local language form, dialect or subdialect you speak/ he/she speaks (best).**		
	1 Kihnu	**4** Setu	**7** Hiiu
	2 Mulgi	**5** Võru	**8** Other (NOTE) ...………..…………....................
	3 Saare	**6** Kodavere	

QUESTIONS A20 AND A20ATO BE ASKED ABOUT PERSONS OF AT LEAST THREE YEARS OF AGE.

A20	**Do you/ does he/she speak other languages besides mother tongue?**
	1 Yes **2** No → **A21**
A20A	**Please name other languages beside mother tongue you speak/ he/she speaks <u>according to the level of language proficiency</u> (begin with the language with the highest level of proficiency).**
	....

QUESTIONS A21 AND A21A TO BE ASKED ABOUT PERSONS OF AT LEAST 15 YEARS OF AGE.

A21	**Do you/ does he/she have any religious affiliation?**
	1 Yes **2** No → **A22** **3** Not willing to answer → **A22**
A21A	**Please indicate your religious affiliation.**
	1 Lutheranism
	2 Orthodoxy
	3 Other (NOTE)………………………………………………………………………………….

Next questions concern your/his/her previous changes of place of residence.

A22 ☼ p 12	QUESTION TO BE ASKED ABOUT PERSONS BORN IN ESTONIA (A13 = 1) AND WITH PLACE OF USUAL RESIDENCE IN THE SAME CITY / RURAL MUNICIPALITY AS THE PLACE OF BIRTH (SEE A06A AND A13A). **Have you/ has he/she continuously lived in the current city / town / rural municipality since birth?** *NB! Continuous residence is not considered interrupted in case of:* *any absence with duration of less than 12 months;* *distant employment, if the respondent stays with his/her household for most of vacant days;* *absence in connection with studies in the general education school (upper secondary school, gymnasium, basic school etc.) or secondary vocational school;* *compulsory military service and participation in a war.* *Studies in an institution of higher education or higher vocational school away from the place of residence are considered to interrupt continuous residence.* **1** Yes → **A28**　　　　　　　　**2** No
A22A **A22AB**	**Indicate the year when you/he/she last settled in this city / town / rural municipality.** Year: |__|__|__|__|
A23	**In which country was your/his/her previous place of residence?** OPTION 2 IS NOTED IF THE PERSON ARRIVED FROM THE MENTIONED TERRITORY TO CURRENT PLACE OF RESIDENCE BEFORE 1945 (A22AB<1945). **1** Estonia **2** Former territory of Estonia – Petserimaa (Pechory) or area to the east of the Narva River → **A28** **3** Russia → **A27** **4** Finland → **A27** **5** Ukraine → **A27** **6** Byelorussia → **A27** **7** Other country (NOTE)…………..…………….…………………→ **A27**
A23A **MK** **V** **A** **Y**	**Please indicate your/his/her previous place of residence in Estonia.** *Consider the place of residence outside city / town / rural municipality where you currently reside/ he/she currently resides. If you are unable to answer on the basis of the current administrative division, indicate the place of residence on the basis of the former administrative division and specify the validity period of that division, as far as you know.* 　　FORMER ADMINISTRATIVE DIVISION County (district) …………………………………………………………………………..………… Rural municipality (village council)…………………….………………………………………………….. City / town / small town / village.. ………………………………………………………………………… |__|__|__|__| Year when the administrative division took effect (in case of former administrative division)
A24 ☼ p 13	QUESTION TO BE ASKED ABOUT PERSONS BORN IN ESTONIA (A13 = 1). **Have you/ has he/she ever lived outside the Republic of Estonia for 12 or more consecutive months?** *Current borders of the Republic of Estonia shall apply.* *NB! Continuous residence outside the Republic of Estonia shall not include:* *employment in a foreign country, if the respondent stays most of vacant days with own household living in Estonia;* *studies in general education school (upper secondary school, gymnasium, basic school etc.) or secondary vocational school in a foreign country if parents' home was in Estonia;* *compulsory military service in the foreign country and participation in a war.* *Studies in an institution of higher education or higher vocational school in a foreign country are considered continuous residence outside the Republic of Estonia.* **1** Yes　　　　　　　　**2** No → **A27A**
A25A **A25B**	**When did you/he/she recently arrive to settle in Estonia? Indicate the year.** *Current borders of the Republic of Estonia shall apply.* Year: |__|__|__|__|

A26	**What was your/his/her most recent country of residence before settling in Estonia?**
	OPTION 2 IS NOTED IF THE PERSON ARRIVED FROM THE MENTIONED TERRITORY TO ESTONIA BEFORE 1945 (A25B < 1945).
	1 Russia
	2 Former territory of Estonia – Petserimaa (Pechory) or area to the east of the Narva River
	3 Finland
	4 Ukraine
	5 Byelorussia
	6 Other country (NOTE)...

QUESTIONS A27 AND A27A TO BE ASKED ABOUT PERSONS BORN BEFORE 31 MARCH 2000, WHO ARRIVED TO THE CURRENT PLACE OF RESIDENCE AFTER 1999 (SEE A22AB).

A27 ☼ p 14	**What was your/his/her country of usual residence on 31 March 2000 (at the time of previous population and housing census)?**
	▪ *In case of a person who was studying at a higher educational institution or post-secondary vocational educational institution at that time, indicate the place of residence based at the place of study.*
	▪ *In case of a person who stayed away from home when studying at a general education school (upper secondary school, gymnasium, basic school etc.) or secondary vocational institution at that time, indicate the place of residence of his/her household.*
	▪ *In case of a conscript, indicate the place of residence before military service.*
	▪ *In case of a person who was residing in an institution (children's home, care home, custodial institution, etc.) for more than a year, indicate the location of the institution.*
	1 Estonia
	2 Russia → **A28**
	3 Finland → **A28**
	4 Ukraine → **A28**
	5 Other country (NOTE)……………………………………………... → **A28**
A27A	**Please indicate your/his/her place of usual residence on 31 March 2000 (at the time of previous population and housing census)?**
	Consider actual place of residence, irrespective of whether it was registered as your/his/her place of residence or not. If you are unable to answer on the basis of the current administrative division, indicate the place of residence on the basis of the former administrative division.
	FORMER ADMINISTRATIVE DIVISION
MK	County……………………………………………………………………………………....................
V	Rural municipality ………………………………………………………………...........................
A	City / town / small town / village..………………………………………………………………

QUESTIONS A28–A47A TO BE ASKED ABOUT PERSONS OF AT LEAST 15 YEARS OF AGE.

Next questions concern education.

A28 ☼ p 15	**What was the level of education you/he/she acquired in general education school (upper secondary school, basic school, primary school etc.)?**
	Please state the highest level of education acquired. As of year 1990 primary education includes 6 grades.
	1 Secondary education → **A29**
	2 Basic education (incl. incomplete secondary education)→ **A29**
	3 Primary education → **A29**
	4 Primary education not acquired
A28A	**Are you/ is he/she…**
	1 literate
	2 illiterate?
	CONTINUE WITH A34

A29 ☼ p 16	**What is your/his/her level of vocational, occupational or professional education?** *Please state your/his/her highest level of education acquired.* **1** Completed higher education → **A33** **2** Completed vocational, occupational or professional education (which is not higher education) **3** No vocational, occupational or professional education → **A34**
A30	**What was the level of education required for entry into the vocational educational institution?** **1** Requirement of secondary education **2** Requirement of basic education → **A31B** **3** No requirement of the level of education → **A31C**
A31A	**In the course of acquiring vocational, occupational or professional education, did you/he/she receive…** **5** secondary specialized education (possible for persons aged 30 or older) **6** other vocational education (incl. vocational secondary education)? **CONTINUE WITH A34**
A31B ☼ p 17	**In the course of acquiring vocational, occupational or professional education, did you/he/she receive…** **1** only vocation **3** vocation and secondary education **4** vocational secondary education (i.e., vocation and a right to be admitted to state examination) **5** secondary specialized education (possible for persons aged 30 or older)? **CONTINUE WITH A34**
A31C	**In the course of acquiring vocational, occupational or professional education, did you/he/she receive…** **1** only vocation **2** vocation and basic education? **CONTINUE WITH A34**
A33 ☼ p 18	**What is your/his/her level of higher education?** *Please state the highest level of education acquired.* **1** Doctor (incl. former Candidate of Science), medical doctor who has completed residency **2** Master (incl. integrated studies of 5 to 6 years, engineering studies, medical doctor who has completed internship, one-year teacher training) **3** Higher education obtained on the basis of a curriculum that was used before 1992 **4** Bachelor (study period of at least 4 years, admission in Estonia until school year 2001/02) **5** Bachelor (study period of 3 years, admission in Estonia from school year 2002/03) **6** Professional higher education, Diploma study (also higher vocational education)

A34 ☼ p 19	**What was your/his/her main source of subsistence in 2011?** **1** Wage and salaries **2** Income from entrepreneurship / farming / business activity **3** Pension for incapacity for work **4** Other pension **5** Maintained by household members or other persons **6** Support, scholarship, benefit (e.g. parental or unemployment insurance benefit, unemployment allowance) **7** Maintained by institution **8** Other source of subsistence

Next questions concern employment during the period 19–25 December 2011 (working week preceding the Census).

A36	**Did you/he/she perform at least one hour of remunerated work during this week (19–25 December 2011)?**
	Include also temporary and unofficial work.
	1 Yes → **A42** **2** No
A37	**Were you/ was he/she temporarily absent from your/his/her place of employment or enterprise during this week (19–25 December 2011) due to leave, illness, etc.?**
	▪ *Temporary absence from work <u>includes</u> also pregnancy and maternity leave (answer "yes").*
	▪ *Temporary absence from work <u>does not include</u> parental leave (answer "no").*
	1 Yes → **A42** **2** No
A38	**Have you/ has he/she been actively seeking work during November 2011?**
	NB! Seeking work also includes any preparations to start business activities or establish a farm, and any waiting periods before the start of agreed employment.
	1 Yes **2** No → **A41**
A39	**If you/he/she would have been offered a suitable job during 19–25 December 2011, could it have been possible for you/him/her to commence work within two weeks?**
	1 Yes **2** No → **A41**
A40	**Have you/ has he/she been employed before?**
	Consider employment for at least three months.
	1 Yes → Indicate the year, when you were/ he/she was employed for the last time. |__|__|__|__| → **A42**
	2 No → **A50**
A41 ☼ p 20	**Which of the following groups did you/he/she belong to during the period 19–25 December 2011?**
	If you consider two groups equally important, indicate the group that comes before the other preference in the list. If you are/he is a conscript, be sure to note "Conscript".
	1 Conscript **5** On parental leave
	2 Student (pupil) **6** Homemaker
	3 Person receiving pension for incapacity for work **7** Unemployed for other reasons
	4 Other pensioner
	CONTINUE WITH A50

FOR EMPLOYED PERSONS (A36 = 1 or A37 = 1). **Next questions concern your/his/her main place of work during the period 19–25 December 2011.**
FOR UNEMPLOYED PERSONS (A40 = 1). **Next questions concern your/his/her last main place of work.**

A42 **A**	**Please state the full name of your/his/her main place of work.**
	………
	………
	If you/he/she worked in a subordinate unit of this institution, then provide subordinate unit's name.
	Subordinate unit is the unit in the structure of the company or institution but having different area of activity or address compared to that of head office.
	………
B	
A43 **A**	**What is (was) the main branch of economic activity of that company/institution?**
	………
	………
	TO BE ASKED IF THE PERSON WORKED IN A SUBORDINATE UNIT.
	What is (was) the main branch of economic activity of that subordinate unit where you/he/she worked?
	………
B	………

A44 A	What is (was) your/his/her occupation at the main place of work?
	………
	Please give a brief description of your/his/her official duties.
B	………
	………

A45	THE QUESTION IS ASKED IF IT CAN BE ASSUMED ON THE BASIS OF OFFICIAL TITLE THAT THE PERSON HAS SUBORDINATES.
	How many employees are (were) directly subordinated to you/him/her?
	Only persons with management responsibilities can have direct subordinates.
	1 No subordinates
	2 1–2
	3 3–25
	4 More than 25

A46 ☼ p 21	**Are (were) you/ is (was) he/she at your/his/her main place of work ...**
	1 employee with a stable contract (incl. public service), whose employment relationship lasts for at least one year
	2 employee with a stable contract, whose employment relationship lasts for less than one year
	3 entrepreneur-employer, farmer with salaried employees
	4 self-employed, farmer without employees, freelancer
	5 other (unpaid worker in a family enterprise or farm; member of commercial association)?
	If you have/he/she has or had multiple statuses at your/his/her main place of work, refer to the place that produces/produced more income.

QUESTIONS A47 AND A47A CONCERN EMPLOYED PERSONS. (FOR UNEMPLOYED PERSONS (A40 = 1) CONTINUE WITH A50.)

A47 ☼ p 22	**Where is the location of your/his/her main place of work?**
	State your / his/her actual location of work. This may differ from the address of the main office of place of work.
	1 At home / own farm → **A50**
	2 Fixed address in Estonia, outside home
	3 No fixed address in Estonia
	4 Fixed address in a foreign country (indicate the country) …………………………………→ **A50**
	5 No fixed address in a foreign country (if you/he/she worked within the territory of one country, please indicate the name of that country.)...→ **A50**

A47A	**Please specify the location of the place of work.**
	IF A47=3 AND PERSON WORKED WITHIN THE TERRITORY OF ONE COUNTY, CITY, RURAL MUNICIPALITY OR SETTLEMENT, PLEASE NOTE THE NAME OF THAT COUNTY, CITY, RURAL MUNICIPALITY OR SETTLEMENT.
MK	County ………………………………………………………………………………………………….............
V	Local government (city / rural municipality) ………………………………………………………………..
A	C. without municipal status / (small) town / village / c. distr.……………………………………...……………..

Next two questions concern your/his/her health.

A50	**Have you/ has he/she any long-term illness or health problem?**
	Do not include any temporary illness or health problem.
	1 Yes **2** No

A51	**What restrictions did health problems place on your/his/her activities of everyday life for longer period (at least for last six months)?**
	1 Significant restrictions
	2 Insignificant restrictions
	3 No restrictions at all
	END OF THE PERSONAL QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED)

QUESTIONS FOR TEMPORARY RESIDENTS

A06	Is your/his/her place of usual residence located in Estonia?
	1 Yes **2** No → **A06B**
A06A ☼ p 5 **MK** **V** **A** **VK** **T** **TL** **M** **K**	**Where is your/ his/her place of usual residence?** County ……………………………………………………………………………………...... Local government (city / rural municipality) …………………………………………………........... C. without municipal status / (small) town / village / c. distr.………………………………………… Small place (e.g., gardening association) ……………………………………………………… Street ………………………………………………………………………………………….... Farm / land unit ………………………………………………………………………………..... House / lot no.…………………………………………………………………………………...... Apartment no.……………………………………………………………...........................……
A07G	**Are any household members present at your/his/her place of usual residence during the Census (January – March 2012) to provide information about your/his/her dwelling and household?** **1** Yes **2** No IF THE ANSWER IS "NO", FILL IN THE HOUSING AND HOUSEHOLD QUESTIONNAIRE AND PERSONAL QUESTIONNAIRES ON SUCH TEMPORARY RESIDENTS. END OF THE PERSONAL QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED)
A06B	**In which country is your/his/her place of usual residence?**
A07A	**How long will you/he/she (probably) stay in Estonia?** **1** Less than 3 months → END OF THE PERSONAL QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED) **2** 3–6 months **3** More than 6 months
A07C	**Name the main reason you are / he/she is staying in Estonia.** 1 Employment 2 Studies 3 Family reasons 4 Other
AV13	**What is your/his/her country of birth?**
AV14	THE QUESTION IS ASKED IF THE PERSON WAS BORN IN ESTONIA. **When did you/he/she last leave to settle abroad?** Year: |__|__|__|__|
AV15	**What is your/his/her country of citizenship?** END OF THE PERSONAL QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED)

QUESTIONS ON PERSONS WHO HAVE LEFT ESTONIA

A06B	In which country is his/her place of usual residence?
AV14	When did he/she last leave to settle abroad? Year: |__|__|__|__|

FINAL CLAUSE: We have reached the end of the personal questionnaire. Thank you for your answers!

ENUMERATOR'S PART (CONTINUED)

Lp	TIME OF COMPLETING INTERVIEW / TIME OF INTERRUPTING INTERVIEW
A	HOUR |__|__|
B	MINUTES |__|__|
Lpk	TIME OF COMPLETING INTERRUPTED INTERVIEW
A	HOUR |__|__|
B	MINUTES |__|__|
KI	LANGUAGE OF INTERVIEW **1** ESTONIAN **2** RUSSIAN **3** OTHER

2011 POPULATION AND HOUSING CENSUS

HOUSING AND HOUSEHOLD QUESTIONNAIRE
(Paper version)

ENUMERATOR'S PART

Kht	LOCATION OF THE INTERVIEW 1 RESPONDENT'S DWELLING 2 RESPONDENT'S PLACE OF WORK 3 ENUMERATOR'S CAR/PLACE OF RESIDENCE 4 OTHER LOCATION (SPECIFY) …………………………………………………………………………
Kpv **A** **B** **C**	DATE OF INTERVIEW **Kpvk** DATE OF CONTINUING INTERRUPTED INTERVIEW DAY |__|__| DAY |__|__| MONTH |__|__| MONTH |__|__| YEAR |__|__|__|__| YEAR |__|__|__|__|
Alg **A** **B**	INTERVIEW STARTED **Algk** TIME OF CONTINUING INTERRUPTED INTERVIEW HOUR |__|__| HOUR |__|__| MINUTES |__|__| MINUTES |__|__|
Kviis	TYPE OF SURVEY 1 – LAPTOP SURVEY 2 – PAPER QUESTIONNAIRE SURVEY
Lik	ENUMERATOR'S PERSONAL IDENTIFICATION CODE I__I__I__I__I__I__I__I__I__I__I__I
Rg	REGIONAL MANAGER NUMBER |__|
Rng	DISTRICT HEAD NUMBER |__|__|
Prk	AREA SUPERVISOR NUMBER |__|__|__|
Jsk	ENUMERATION AREA NUMBER |__|__|__|__|
Larv	NUMBER OF HOUSEHOLDS IN DWELLING (Question E04) |__|__|
Lnr	SEQUENTIAL NO. OF HOUSEHOLD IN DWELLING I__I__I *Household questionnaire document no. in FMS I__I__I__I__I__I* **to be filled in by area supervisor** *Dwelling questionnaire document no. in FMS I__I__I__I__I__I* **to be filled in by area supervisor**
Iarv	NUMBER OF PERSONAL QUESTIONNAIRES IN THE SET (List A + List B) I__I__I
Aonr	DATA OBJECT NO. I__I__I__I__I__I__I__I__I__I__I__I__I **to be filled in by enumerator**

Tallinn 2011

HOUSING AND HOUSEHOLD

At first please provide information about dwelling.

E01	**Address of the dwelling**
MK	County ……………………………………………………………………………………..................
V	Local government (city / rural municipality) ……………………………………………………..................
A	C. without municipal status / (small) town / village / c. distr.…………………………………......…………………
VK	Small place (e.g., gardening association) …………………………………………………………..........
T	Street …………………………………………………………………………………………….........…
TL	Farm / land unit …………………………………………………………………………………......……
M	House / lot no.……………………………………………………………………………………….........
K	Apartment no.……………………………………………………………………………………....................

Eas	**LOCATION OF DWELLING**
	1 DWELLING IN THE ENUMERATOR'S ENUMERATION AREA → **E02**
	2 DWELLING LOCATED ELSEWHERE, WITH PERMANENT RESIDENTS (USUAL PLACE OF RESIDENCE OF A TEMPORARY RESIDENT)

| E01L | **PRECISE ADDRESS OF THE LOCATION OF THE CENSUS INTERVIEW:**……………………………………….....|
| | …….....|

QUESTIONS E02–E04 ARE FILLED IN ABOUT THE ENTIRE DWELLING, IRRESPECTIVE OF WHETHER THE HOUSEHOLD USES ENTIRE OR PART OF THAT DWELLING OR THE DWELLING IS EMPTY. FOR EXAMPLE, IF A HOUSHOLD RENTS ONE ROOM OF AN APARTMENT, THEN ANSWER "APARTMENT…" TO QUESTION E02.

E02 ☼ p 1	**What is the type of the dwelling?**
	1 Apartment in an apartment building
	2 Apartment in non-residential building
	3 Private house (one-family dwelling)
	4 Apartment in a private house divided into several dwellings (e.g. in a house that was originally built as one-family dwelling)
	5 Box of the semi-detached house
	6 Box of the terraced house
	7 Room (rooms) in a hostel or in an accommodation establishment → **E04**
	8 Social welfare institution providing temporary accommodation (e.g. shelter, social housing unit) → **E04**
	9 Establishment, institution (e.g. children's home, youth home, care home, custodial institution, monastery, etc.) → END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE
	10 Other room or a building (e.g. summer cottage not suitable for year-round living; trailer; boat; garage) → **E04**
	11 Roofless (homeless) → **EX2**
	12 I am a permanent resident of Estonia, but I do not have a place of residence in Estonia at the moment, as I am residing temporarily abroad (e.g. a diplomat) → **EX2**

E03 ☼ p 2	**Which is the occupancy of the dwelling?**
	1 At least one resident of Estonia is permanently living in the dwelling
	2 Only persons not subject to enumeration are living in the dwelling (e.g., foreign diplomats and their family members; foreign residents staying in Estonia for less than three months) → **EX2**
	3 There are no permanent residents in the dwelling (may be temporary residents)→ **EX2**

E04	How many households permanently reside in this dwelling?
	Household is a group of people usually living in a common dwelling, who share available household facilities (common budget and food), also a person living alone is also a household. A household is similar to a family but, unlike a family, it could also have only one member or include non-relatives.
	Number of households: |__|__|

HOUSEHOLD

IF THERE ARE SEVERAL HOUSEHOLDS LIVING IN ONE DWELLING, TAKE AN EMPTY HOUSING AND HOUSEHOLD QUESTIONNAIRE FORM FOR EACH FOLLOWING HOUSEHOLD, WRITE DOWN THE NUMBER OF THE HOUSEHOLD IN THE DWELLING (Lnr) AND FILL IN THE QUESTIONNAIRE TOGETHER WITH PERSONAL QUESTIONNAIRES AFTER YOU HAVE FINISHED INTERVIEWING THE CURRENT HOUSEHOLD.

EX2	**WHO SHALL PROVIDE ANSWERS FOR THE QUESTIONNAIRE?**
	1 MEMBER OF THE HOUSEHOLD IN QUESTION
	2 PERSON LIVING IN THE DWELLING FOR A SHORT PERIOD
	3 OTHER RELIABLE PERSON
	4 ENUMERATOR → **E15**
EL05 A B	**Please provide your contact information in order to allow specification of information.**
	Telephone ……………………………...…
	E-mail …………………………………...…
EX3	**DIRECTION TO NEXT QUESTION:**
	If E04 > 0 or E02 = (11 or 12) → **List A**
	If (E03 = 3 and EX2 = 2) or E04 = 0 → **EL07** (temporary residents)
	In other cases, if E03 ≠ 1 → **E15** (questions on dwelling)

HOUSEHOLD MEMBERS (List A)
☼ p 3

Please indicate all members of your household (including yourself), who were <u>permanent</u> residents at this address as of 31 December 2011. Include also these members of your household who were temporarily (less than 12 months) absent.

- *Include in your household also:*
    - *a person who is working away from home for a period exceeding 12 months, but stays with your household for most of his/her days off;*
    - *a pupil of secondary vocational educational institution or general education school (gymnasium, basic school) residing away from home during studies;*
    - *a conscript.*
- *Do not include in your household:*
    - *a student (pupil) of a higher educational or post-secondary vocational educational institution, residing elsewhere during studies;*
    - *a person who has been or is probably residing in an institution (e.g. care home) for a period exceeding 12 months.*

No.	Given name and surname (USE BLOCK LETTERS)	Personal identification code	Sex (M–1 F–2)	Date of birth (day, month, year)	Age (31.12. 2011)
1.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
2.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
3.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
4.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
5.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
6.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
7.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
8.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
9.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
10.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
11.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	
12.		|_|_|_|_|_|_|_|_|_|_|		|_|_| |_|_| |_|_|_|_|	

READ OUT THE NUMBER AND NAMES OF HOUSEHOLD MEMBERS.

EL06	**Did you leave someone out because you are not sure whether he/she belongs to your household?** *For instance person who also has another place of residence or person temporarily absent.* **1** Yes → IF NECESSARY, ENTER THE PERSON IN THE LIST **2** No
EX5	**DIRECTION TO NEXT QUESTION:** If dwelling is missing (E02 = 11 or 12) and there is one person on List A → END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE, INTERVIEWER'S PART (CONTINUED) If dwelling is missing (E02 = 11 or 12) and there are several persons on List A → **EL13**

To enumerate all people, we also ask information on temporary residents.

EL07
☼
p 4
	Are there temporary residents who live in this dwelling during some time in a week or year (e.g., former household members)?
	
	A temporary resident is someone whose place of usual residence is elsewhere.
	
	Please include:
	
	   ▪ *a resident of Estonia who stays in this dwelling for at least three months (90 days) per year (e.g., a university student who studies elsewhere). Non-consecutive days count as well. For example, if a person stays in the dwelling on all weekends of the year, it makes more than 100 days in total;*
	
	   ▪ *a resident of Estonia who does not stay in his/her place of usual residence for most of the census period (31 December 2011 – 31 March 2012);*
	
	   ▪ *a resident of a foreign country who arrived in Estonia before 31 December 2011 and intends to stay in Estonia for at least three months (90 days).*
	
	1 Yes **2** No

EX6	**DIRECTION TO NEXT QUESTION:** If List A is empty and EL07 = 1 → **List B** If List A is empty and EL07 ≠ 1 → **EX11**

EL08	**Has any close relative of yourself or of a member of your household (e.g., parent, child, sister, brother, spouse/cohabitant) left Estonia in 2000 or later and is currently living abroad?**
	
	   ▪ *Consider also students in universities or post-secondary vocational education institutions who have lived abroad for 12 months or more due to their studies.*
	
	   ▪ *If you have previously listed a close relative, who had left Estonia, as a temporary resident, do not include such relatives here.*
	
	1 Yes **2** No

EX7	**DIRECTION TO NEXT QUESTION:** If EL07 ≠ 1 and EL08 ≠ 1 → **EX11**

TEMPORARY RESIDENTS AND PERSONS WHO HAVE LEFT TO LIVE ABROAD (List B)

Please state the details of temporary residents in the dwelling and/or the details of close relatives who have left Estonia.

	No. of the person	B1	B2	B3	B4
A01	**Given name**				
A02	**Surname**				
A03	**Personal identification code**				
A04	**Sex** **1** Male **2** Female	|＿|	|＿|	|＿|	|＿|
A05 A B C D	**Date of birth** Day Month Year Age as of 31 December 2011	|＿|＿| |＿|＿| |＿|＿|＿|＿| |＿|＿|＿|	|＿|＿| |＿|＿| |＿|＿|＿|＿| |＿|＿|＿|	|＿|＿| |＿|＿| |＿|＿|＿|＿| |＿|＿|＿|	|＿|＿| |＿|＿| |＿|＿|＿|＿| |＿|＿|＿|
AX1	**Person's status** **2** Temporary resident **3** Permanently left Estonia	|＿|	|＿|	|＿|	|＿|
	[CONTINUE A01 UNTIL ALL PERSONS HAVE BEEN ENTERED IN THE LIST]				

READ OUT THE NAMES OF THE PERSONS INCLUDED ON THE LIST AND THEIR STATUSES.

EX11	DIRECTION TO NEXT QUESTION: If lists A and B include only one person in total or both lists are empty → **EX12**

Next questions concern relationship between persons.

THE NAMES OF PERSONS SHALL BE TAKEN FROM LIST **A** AT FIRST, THEN FROM LIST **B**.
RELATIONSHIPS HAVE TO BE IDENTIFIED BY ROWS. AT FIRST WRITE DOWN THE RELATIONSHIP OF THE SECOND PERSON WITH RESPECT TO THE FIRST PERSON, THEN THE RELATIONSHIP OF THE THIRD PERSON WITH RESPECT TO THE FIRST PERSON AND THE SECOND PERSON, AND SO ON.

If several options are suitable for noting relationship, select the first option.

EL13 ☼ p 6	Who is /the 2nd person/ with respect to /the 1st person/? etc.		1	2	3	4	5	6	7	8	9	10	11
		1	■	■	■	■	■	■	■	■	■	■	■
	1 Spouse	2		■	■	■	■	■	■	■	■	■	■
	2 Cohabitant **3** Child (incl. adopted)	3			■	■	■	■	■	■	■	■	■
	4 Child of the spouse or cohabitant	4				■	■	■	■	■	■	■	■
	5 Mother/father (incl. adoptive parent)	5					■	■	■	■	■	■	■
	6 Spouse or cohabitant of mother/father	6						■	■	■	■	■	■
	7 Sister/brother (incl. half-sister/half-brother or child of the parent's spouse/cohabitant)	7							■	■	■	■	■
	8 Grandparent (incl. spouse/cohabitant of a grandparent)	8								■	■	■	■
	9 Grandchild (incl. grandchild of the spouse/cohabitant)	9									■	■	■
	10 Other relative (incl. spouse's/cohabitant's relative)	10										■	■
	11 Not related	11											■
		12											

EX12	DIRECTION TO NEXT QUESTION: If List A is empty and E02 = (1–6) or 10 → **E15** If List A is empty and E02 = (7 or 8) → END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED) If dwelling is missing (E02 = 11 or 12) → END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED)

EL14	ASK THE QUESTION IF E02 = (1–6) or 10. **What is the basis <u>for your household</u> to use this dwelling?** **1** Dwelling belongs to the household (a member of household) **2** Household rents the dwelling **3** Free use (household does not pay rent to the owner, but may pay for utility services)

EL27	**Does your household own a dwelling with no permanent residents during the Census (January – 31 March 2012)?**
	Consider only apartments, private houses, terraced houses or semi-detached houses on the territory of Estonia.
	Do not take into account cottages which are not suitable for all-year-round habitation, and the premises that are used for business or professional purposes only.
	1 Yes **2** No → **EL28**
EL27A	**How many such dwellings does your household own?** |__|__|
EL28 ☼ p 7	**Does your household grow agricultural and horticultural products or keep farm animals, domestic fowls or bees mainly for own consumption?**
	"Yes" should always be noted if the area of land used for growing agricultural crops or horticultural products or keeping farm animals is <u>less than one hectare</u> (10,000 m²) and the household has at least one of the listed options:
	– 50 m² of land for growing vegetables or
	– three fruit trees or
	– six berry bushes or
	– 10 rabbits or
	– one other farm animal (pig, bovine animal, sheep, goat) or
	– 10 domestic fowls or
	– three beehives.
	1 Yes **2** No
EX13	**DIRECTION TO NEXT QUESTION:**
	If E02 = (7 or 8) and EL27= 1 → PART "VACANT DWELLINGS"
	If E02 = (7 or 8) and EL27= 2 → END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED)

DWELLING

QUESTIONS E15–E26 ARE FILLED IN ABOUT THE ENTIRE DWELLING, IRRESPECTIVE OF WHETHER THE HOUSEHOLD USES ENTIRE OR PART OF THAT DWELLING OR THE DWELLING IS EMPTY.

Next questions concern your dwelling. (ADDRESS E01)

E15 ☼ p 8	**Who is the owner of this dwelling?**
	If the dwelling has multiple owners, note the option that is higher on the list.
	1 Resident(s) of Estonia (place of usual residence is in Estonia)
	2 Resident(s) of foreign country (place of usual residence abroad)
	3 State or local government
	4 Other (company, non-profit association, legal person governed by public law etc.)
E16 ☼ p 9	**When was this house built?**
	Consider the period which includes the year when the house was officially approved.
	1 Before year 1919 **5** 1971–1980 **9** In 2006 or later **11** Exact time unknown, but before year 1946
	2 1919–1945 **6** 1981–1990 **10** Uncompleted building **12** Exact time unknown, but during 1946–1990
	3 1946–1960 **7** 1991–2000 **13** Exact time unknown, but in 1991 or later
	4 1961–1970 **8** 2001–2005
E17	**How many rooms does the dwelling have?** |__|__|
E18	**What is the total area of the dwelling?** |__|__|__|__| m²
	(Useful) floor area of the dwelling is the total aggregated floor area of rooms, kitchen and ancillary premises inside the dwelling.

E19	**Is there a kitchen in the dwelling?**
	1 There is a separate kitchen
	2 Kitchen forms a part of the room or there is a kitchenette
	3 The dwelling has no kitchen or kitchenette
E21	**Is the dwelling equipped with water supply (is there an opportunity to use running water)?**
	1 Yes **2** No → **E24**
E23	**Does the dwelling have a bath or a shower?**
	1 Yes → **E25** **2** No
E24	**Does a sauna belong to the dwelling?**
	1 Yes **2** No
E25	**Is the dwelling equipped with a toilet?**
	1 The dwelling has flush toilet
	2 The dwelling has dry toilet
	3 The dwelling has no flush toilet or dry toilet
E26 ☼ p 10	**What is the main heating option used in the dwelling?**
	1 Distant central heating
	2 Local central heating (with a boiler or heat pump)
	3 Stove or fireplace heating
	4 Electrical heating (e.g. electric or oil heaters, electrical floor heating)
	5 Heating options are not available
EX14	**DIRECTION TO NEXT QUESTION:** If List A is empty or EL27 = 2 → END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE, ENUMERATOR'S PART (CONTINUED)

VACANT DWELLINGS

You stated before that your household owns dwelling(s) with no permanent residents during the Census. Please provide the following information about this/these dwelling(s).

THE FOLLOWING QUESTIONS SHALL BE ANSWERED IN VIEW OF THE ENTIRE DWELLING, IRRESPECTIVE WHETHER THE HOUSEHOLD OWNS THE ENTIRE DWELLING OR NOT.

	Sequential number of dwelling	1	2	3	4	5
E01	**Where is this dwelling located?**					
MK	County	………………	………………	………………	………………	………………
V	Local government (city / rural municipality)	………………	………………	………………	………………	………………
A	C. without municipal status / (small) town / village / c. distr	………………	………………	………………	………………	………………
VK	Small place (e.g., gardening association)	………………	………………	………………	………………	………………
T	Street	………………	………………	………………	………………	………………
TL	Farm / land unit	………………	………………	………………	………………	………………
M	House / lot no	………………	………………	………………	………………	………………
K	Apartment no	………………	………………	………………	………………	………………

E02 ☼ p 1	**What is the type of the dwelling?** **1** Apartment in an apartment building **2** Apartment in non-residential building **3** Private house (one-family dwelling) **4** Apartment in a private house divided into several dwellings (e.g. in a house that was originally built as one-family dwelling) **5** Box of the semi-detached house **6** Box of the terraced house	|＿|	|＿|	|＿|	|＿|	|＿|
E16 ☼ p 9	**When was this house built?** *Consider the period which includes the year when the house was officially approved.* **1** Before year 1919 **2** 1919–1945 **3** 1946–1960 **4** 1961–1970 **5** 1971–1980 **6** 1981–1990 **7** 1991–2000 **8** 2001–2005 **9** In 2006 or later **10** Uncompleted building **11** Exact time unknown, but before year 1946 **12** Exact time unknown, but during 1946–1990 **13** Exact time unknown, but in 1991 or later	|＿|＿|	|＿|＿|	|＿|＿|	|＿|＿|	|＿|＿|
E17	**How many rooms does the dwelling have?**	|＿|＿|	|＿|＿|	|＿|＿|	|＿|＿|	|＿|＿|
E18	**What is the total area of the dwelling?** *(Useful) floor area of the dwelling is the total aggregated floor area of rooms, kitchen and ancillary premises inside the dwelling.*	|＿|＿|＿|	|＿|＿|＿|	|＿|＿|＿|	|＿|＿|＿|	|＿|＿|＿|
E19	**Is there a kitchen in the dwelling?** 1 There is a separate kitchen 2 Kitchen forms a part of the room or there is a kitchenette 3 The dwelling has no kitchen or kitchenette	|＿|	|＿|	|＿|	|＿|	|＿|
E21	**Is the dwelling equipped with water supply (is there an opportunity to use running water)?** **1** Yes **2** No → **E24**	|＿|	|＿|	|＿|	|＿|	|＿|
E23	**Does the dwelling have a bath or a shower?** **1** Yes → **E25** **2** No	|＿|	|＿|	|＿|	|＿|	|＿|

E24	**Does a sauna belong to the dwelling?** **1** Yes **2** No	|__|	|__|	|__|	|__|	|__|
E25	**Is the dwelling equipped with a toilet?** **1** The dwelling has flush toilet **2** The dwelling has dry toilet **3** The dwelling has no flush toilet or dry toilet	|__|	|__|	|__|	|__|	|__|
E26 ☼ p 10	**What is the main heating option used in the dwelling?** **1** Distant central heating **2** Local central heating (with a boiler or heat pump) **3** Stove or fireplace heating **4** Electrical heating (e.g. electric or oil heaters, electrical floor heating) **5** Heating options are not available	|__|	|__|	|__|	|__|	|__|
	[CONTINUE WITH E01 UNTIL INFORMATION ABOUT ALL DWELLINGS HAS BEEN ENTERED IN THE LIST]					

END OF THE HOUSING AND HOUSEHOLD QUESTIONNAIRE

ENUMERATOR'S PART (CONTINUED)

Lp A B	TIME OF COMPLETING INTERVIEW / TIME OF INTERRUPTING INTERVIEW HOUR |__|__| MINUTES |__|__|
Lpk A B	TIME OF COMPLETING INTERRUPTED INTERVIEW HOUR |__|__| MINUTES |__|__|
KI	LANGUAGE OF INTERVIEW **1** ESTONIAN **2** RUSSIAN **3** OTHER