**Cooperation with Scientific**

**Communities**

Diana Beltadze, Statistics Estonia, diana.beltadze@stat.eee

**Abstract**

*The presentation introduces practices for developing cooperation with the scientific community in the area of changes in census methodology.*

*Statistics Estonia has decided to organise the next census as a register-based census, and a methodology based on indices has been developed for this. Statistics Estonia chose a cooperation strategy which involves explaining the benefits of the new methodology relative to the objectives of the target group and building cooperation with groups of researchers.*

*The key to success depends on how well the partners know the values of the organisation and what is the capability of Statistics Estonia to communicate the census methodology objectives in an understandable way.*

**Keywords**:census, cooperation with scientific communities, background for partnerships

## Introduction

## In 2009, Statistics Estonia decided to start preparing for a register-based census in the next census round. The decision was made at the time when the first combined-method census of 2011 was being organised. There was little staff as well as expertise available for a register-based census. The solution was to invite scientists and data experts from different institutions to cooperate.

## Why was it necessary to plan this work during such a hectic time? The most important reason lies in today’s rapidly developing society; in a situation where there is a risk of economic and political turbulences, the traditional census is not up-to-date anymore, as it is carried out every 10 years and data analysis can take more than a year. Another reason why new methodology of census is necessary is the increasing mobility of population and the people’s unwillingness to disclose their personal data. It must be possible for the state to monitor its most important resource, i.e. population, continuously and estimate it comprehensively and quickly. This is best achieved through register-based censuses. Therefore, one of the key pillars of the country’s strategic development is to ensure capacity to organise register-based censuses as often as needed, also yearly, if there arises need for it.

## In order to achieve this objective, Statistics Estonia will have to update the IT solution for statistics production, so that it would allow for a process that is as automated as possible. In addition, the solutions must be sustainable in the longer term and support an increase in the production of current statistics on the basis of registers. To this end, it was necessary to provide for necessary development work, and more importantly, cooperation between different authorities, in order to ensure that the data can be used accordingly.

## In the context of conducting censuses, it should be noted that cooperation with scientific communities’ authorities’ fosters partnerships with authorities and experts of different domains. Through partnerships, it has been possible to work together to develop, prepare and implement common solutions to shared problems and challenges that are specific to our country. Also, Statistics Estonia has created an information environment and engage in public advocacy work to explain what is done and why, which are the methodological innovations in statistics presented.

## Cooperation with other authorities was sought in particular with regard to the fact that the quality of the register-based census is determined by the quality of the register data. A high quality of data in the registers ensures high quality of census data. The key issue for Statistics Estonia was to develop a suitable framework for cooperation with different authorities answering for the content and quality of registers.

## Today, it can be said that the cooperation resulting from the preparation of the register-based census contributes significantly to the development of Estonia as an e-state, as the registry system and the data transporting environment which have been organised and coordinated as well as tested and improved through pilot censuses has very high value in its own right. Because of that it will be possible to cut red tape significantly in each sphere of life in the country making data handling more effective and transparent. In addition, new opportunities involving the use of big data allow more operational monitoring of different processes essential in population statistics and other areas of official statistics.

## Successive register-based censuses are highly comparable, as they generally rely on the same methodology and dataset. This makes it possible to assess very accurately population trends in our small but heterogeneous country.

## Cooperation with the parties involved is very diverse and has many forms, but this work contributes to the implementation of the idea underlying the institution’s long-term strategic activity – organisation of a register-based census in 2021. Cooperation with research institutions is key to the success of this strategy. Investing in this requires recognising that other associations and organisations are not competitors but partners, and that supporting each other is important.

During long time, Statistics Estonia has had very good partnership with the University of Tartu (UT), especially with the institute of mathematics and statistics. The academic personal of this institute has participated in solving problems connected with population statistics and especially with census (Tiit, Meres, Vähi, 2012), (estimation of under-coverage of census), also some prosperous ideas (assessing the adequate population size, estimation the household members status) have been evolved in master theses defended in this institute (Maasing, 2015), (Kütt, 2014).

# Some other institutes of UT have also been long-term cooperation partners for Statistics Estonia. In population studies is the Institute of Ecology and Earth Sciences very good partner, participating in the studies connected with population projection and migration. The most remarkable contribution from their side is the common project using the big data from cellular phones for geographical positioning. Also with colleagues from the institute of social studies of UT Statistics Estonia has very fruitful cooperation, the surveys collected in cooperation with them are fruitful testing material for methodologies developed for census.

Another active academic partner for Statistics Estonia is Tallinn University (TLU), School of Governance, Law and Society. The scientists of this unit help statisticians as opponents of methodologies elaborated in Statistics Estonia.

# 

## A key issue in preparing for the register-based census has been with whom and how cooperation should take place, in order to better achieve the objectives and foresee what will influence the forms of cooperation. The latter is in particular linked to Statistics Estonia’s own call to notice the needs and interests of other involved parties, as good cooperation should be guided by the strategic interests of both organisations. The census team’s intention has been to focus rather on substantive cooperation, where both sides will contribute to the achievement of the set objectives by different means.

## The bigger cooperation projects have been financed from the census budget. Consultation activities have taken place as voluntary work carried out by researchers.

## In this article, the forms of cooperation with research institutions, are looked at in particular with regard to the methodology of the register-based census. The effectiveness of these in meeting the institution’s objectives are assessed. The article highlights fruitful cooperation projects, e.g. in the field of data mining and evaluation of registers, provides recommendations to address possible cooperation aspects on census topics with the broader scientific community and discusses to a significant extent the role of users of census data in the collaborative research framework.

## Initiative and motivation for cooperation

Before the last traditional census, in 2008, in Statistics Estonia was created a Research Council of census and register-based census, where alongside statisticians also scientists from Tallinn and Tartu Universities and some other academic institutions discussed topic problems of census, register-based census and connected themes.

Statistics Estonia started preparations for the register-based census in 2008 by mapping the needs of the register-based census. Inviting scientists to co-operate started in 2009 when the project for working out methodology for a register-based population and housing census was launched.

At the same time, Statistics Estonia was preparing a census based on combined census methodology in 2011; therefore, extra forces were needed for working on longer-term census plans.

The new project was carried out by the consulting firm Ernst &Young Baltic AS and the Estonian Institute for Population Studies from the University of Tallinn, in collaboration with researchers, academic staff and doctoral students from the Institute of International Social Studies from the University of Tallinn, Institute for Social Work from the University of Tallinn, Institute for Ecology and Earth Sciences from the University of Tartu, School of Economics and Business Administration from the University of Tartu, and Institute of Computer Science of the University of Tartu.(Puur, Sakkeus, Aben, 2013).

Such a project for the inclusion of researchers was not uncommon (one of the first projects targeted at cooperation with scientists had been launched already in 1996, which also focused on the development of register-based census methodology), but this methodology project was large scale in time scale as well as in terms of the number of topics and persons involved in comparison with previous scientific co-operation projects.

The aim of the project was to assess the readiness for the register-based population and housing census, to analyse the status of the registers for the census requirements, and develop an initial methodology to enable conducting a register-based pilot census in 2014 (Lehto, 2014).

In the project preparation phase, the results of previous registry analyses conducted or commissioned by Statistics Estonia in the period 2007-2011 were worked through. Additionally, we studied the methodology and experiences of register-based censuses in other countries.

First, the requirements for obligatory characteristics laid down in the regulations of the Council of Europe and the European Parliament, as well as the regulations of the European Commission were analysed, the coverage of each census characteristic was mapped and suggestions were made for the formation of census characteristics and for quality analysis (EU 763, 2008).

Secondly, the practical set-up of the characteristics was tested on the basis of individual datasets and the quality (coverage and accuracy) of characteristics formed on the basis of the registers was assessed. Experts from various scientific fields were involved in the project, who, in cooperation with the employees of Statistics Estonia prepared materials for the final report. The state and quality of the registers of that time (2010—2013) were thoroughly analysed and on the basis of these, a conclusion was drawn that it is not possible to perform a register-based census at the required level.

Cooperation was very effective. The main value of the results was the systematic mapping of the data in the registers. The results of the project were publicly presented in 2013. At the same time, Statistics Estonia, in cooperation with the registers, had started work on improving the data quality of registers and methodological work on implementing systematic cross-usage of registers to ensure adequate information. Despite the critical opinions of the authors, Statistics Estonia continued preparations for the register-based census.

The target for 2014 was to work out a package of legal and organisational measures to improve the quality, timeliness and coverage of the dataset for the register-based census based on the bottlenecks pointed out in the methodological report.

Cooperation with scientists from various fields continued in the Research Council; after it terminated its activities in 2016, in numerous workshops and focus groups, where representatives of various research institutions reviewed and discussed the methodological solutions developed by Statistics Estonia for the formation of census characteristics and assessment of the results.

The key activity after the methodological report was published was to increase awareness of the public, during which the requirements of the register-based census, in particular the availability and quality of data, were presented. A lot of attention was given to the significance of data adequacy, accuracy and operational upgrading in the databases. Statistics Estonia suggested to ministries to make relevant changes to laws to ensure the acquisition of high-quality data by Statistics Estonia. In cooperation with the Estonian Information Systems Authority, a project was initiated, during which the law firm LMP carried out an analysis of the basic regulation of the databases and worked out legislative amendments that would ensure that the register-based census is carried out more efficiently.

At the end of 2014, Statistics Estonia presented the results of the pilot census as the first significant work in the preparation for the register-based population and housing census. The concept of the register-based census and the related software systems were also tested in the process on this work.

By this time the original index-based methodology for current assessment of the population size was elaborated by the team of researchers from UT and statistics Estonia (Maasing, Tiit, Vähi, 2017), where a big number of administrative registers was used to warrant high quality of estimation. The results were tested against surveys and the estimation error was less than 1%.

In conclusion, it should be noted that synergies were achieved in cooperation with scientists, a long-term project was launched for the register-based census. Giving feedback on the state of registers by scientists and statisticians was a driving force for the systematic development of state registers.

Cooperation with the scientific community will continue in the same form as after the end of the methodology project: the methodological steps taken by Statistics Estonia will be presented to various scientific communities in Estonia and abroad, while the methodologists of Statistics Estonia take full account of the critical remarks, but do not give up the preparation of the register-based census, being convinced that at the current level of the Estonian registers, it would be possible to achieve the required accuracy of the census results (in some characteristics these may be in the form of statistical estimates, which is an internationally acceptable methodology).

Cooperation with the partners was smooth and valuable. This cooperation was definitely ambitious and effective also for Statistics Estonia. Bottlenecks in communication with a group of representatives of TLU were about whether to start preparations for the transition to register-based census or accept and wait for natural developments of the registries and progress in the quality of data over time. Statistics Estonia does not consider it expedient to apply the survey methodology in the next census round (not even for individual questions) because the results thus obtained are not sustainable in the context of register-based statistics.

## Cooperation reality: both support and criticism are useful

It is natural to use national databases for statistical purposes in a register-based census. The use of registers in Estonia is characterised by two specific aspects: first, the system of registers is well designed and interoperable; second, it should be taken into account that the registration culture is low and for some reasons, erroneous data are often entered into registers. This concerns particularly the residence data: the change of residence was often not registered in the population register. If the person was leaving the country, they wanted to keep various benefits offered to Estonian residents by formally retaining their previous place of residence (Äär, 2015). As a result, the population number in the population register, which also included thousands of people who had left Estonia, was inaccurate. To tackle this problem, the index-based methodology was worked out on the basis of several national registries for determining the population in the current year. The methodology was tested during a number of years, and in 2015, Statistics Estonia started to produce index-based statistics (Tiit, Maasing, 2016), From the same year the index-based methodology was also applied for estimation the migration that allows to count not only registered, but also non-registered migration (Tiit, Maasing, Visk, Vähi, 2018).

Statisticians are often criticised because of the methodological changes. This also includes Statistics Estonia, when after publishing the number of population in 2016 it turned out that the population that had been declining so far, increased for the first time (Statistics Estonia, 2016). However, the main reason for the population growth was not the change in methodology, but the increased immigration from other European Union countries. Both immigration and emigration increased, as the new methodology allowed also to take into account unregistered migration. It is now clear that Estonia’s population is growing indeed due to the impact of immigration exceeding emigration, whereas the immigrants include both returnees and immigrants arriving from abroad. However, the relatively large migration turnover due to the methodology, with a simultaneously good balance between immigration and emigration, causes questions to migration investigators. It is possible that the index-based calculation of immigration and emigration is influenced by the relatively large commuting in Estonia.

The original index-based methodology for the calculation of population developed at Statistics Estonia is simple and logical, based on the mutual control of the registry data by their cross-usage (Tiit, Vähi, 2017). The number of population estimated with this method has been tested with research data and it is evidently much more precise than the number of population obtained in the last two censuses (under-covered) and the number of population in the population register. However, the intention to implement it in the register-based census has caused opposition in some government authorities. Statistics Estonia had to protect the method and explain it to those authorities that had criticized the method. These were some government departments, the census committee of the Estonian Government, members of the parliamentary committee of demographic crisis.

The most critical part of the community is researchers who are active users of the census results and who have very high expectations towards census results.

In the last census, in 2011, the census questionnaire in Estonia was rather long. Besides questions common for all EU members, in Estonian census were asked ethnic nationality, command of state language and all other foreign languages, and also local dialects, religion, parents’ and grandparents’ place of birth, some questions on health (Metoodika, 2014). Argumentation for that was that this was probably the last traditional census.

The researchers who used this information in 2011 are very much interested to get the information in 2011 again to analyse the changes. But in this list there are questions that cannot be answered by the information got from administrative registers (all foreign languages and dialects, self-estimation of health status etc). That is one reason why some researchers are in favour of traditional census instead of register-based census.

One problematic question was also to estimate the distribution of household types. The second register-based pilot census (2016) gave the distribution that substantially differed from the households’ distribution of census 2011, especially by the share of lone parents. The reason for it is, seemingly, massive wrong registration of living-room due to different possible bonuses. From here a new task arose: to elaborate methodology for finding partners and forming family nuclei using data from administrative registers. This task was solved (Visk, 2018) and the results are quite promising.

The cooperation with scientific communities took the form of focus group meetings in a narrow circle, where there was always an open and fair exchange of information. This form of work has contributed to the clarification of the census method and provoked thinking. Scientists from the Universities of Tallinn and Tartu have been involved; they are really interested in the development of the census methodology. The focus group members have become our “critical friends” in the census methodology, they read the materials and then give us feedback. The interviews take place once a quarter, where, in addition to the methodological updates, we also explain the context of the general preparatory work for the register-based census.

Statistics Estonia also has cooperation projects launched by researchers from which we hope to benefit substantially. One of these is the methodology for using mobile data, which is being developed by scientists from the University of Tartu. This information will enable us to confirm the person’s residency and clarify his place of residence. Another interesting problem relates to data on electricity consumption, which is investigated by researchers of Tallinn University of Technology — these data make it possible to verify housing data, which also are part of the census.

The attitude of researchers toward the new population calculation and census methodology is not the same. Cooperation and openness have not diminished the skepticism of some researchers regarding the correctness of the methodology for the register-based census, who believe the register-based census in 2021 to be too early (Memorandum, 2018). By contrast, there are scientists whose attitudes towards the introduction of the methodology are encouraging and who are enthusiastic about what we are doing to introduce the new technologies. Among them, there are good co-thinkers and advisers, to whom we are very grateful. We know that and respect different points of view.

To sum up, there are several trends in cooperation with the scientific community: one of them focuses on solving the problem of data quality by using the old data collection methods. The second offers modern, very bold information technology solutions for solving the problems. In cooperation with various representatives, it is important that the concept used by Statistics Estonia is based on the responsibility for the quality of the census outcomes. In this regard, we are ready to collaborate with various research institutions to explore different data and to find ways to introduce new data sources and methodologies for producing statistics.

## Conclusions

The cooperation of Statistics Estonia with the scientific community in the census topics has been important and necessary. The best example is cooperation with the Institute of Mathematics and Statistics of UT, where members of staff have participated in solving methodological problems connected with census, and some master projects also are dedicated to census problems.

We have offered cooperation to scientists from various research institutions by informing them of all methodological work, first in the Research Council, later in workshops and in focus groups. The form of collaboration is the reviews and critical remarks made by the researchers, which we have taken into account (as far as possible). The factors contributing to this collaborative process have always been linked to the shared values of the organisation’s representatives.

This presentation provided an overview of cooperation with research institutions in the preparation of the register-based census. It is very positive to note that, as regards all forms of cooperation, researchers have been willing to cooperate, which is directly related to the expected outcome of the census. In most cases, it can be noted that our communication with the universities has always involved an expectation of receiving support or assistance to finding solutions. The motivation of Statistics Estonia in developing the methodology has been to have a sustainable solution to the many-years-long place of residence problem and in the long term, to create opportunities for the state to save costs and make the production of statistics more efficient.

However, when planning cooperation, it has to be taken into account that the motivation of the researchers to participate in the cooperation may be multifaceted. On the one hand, there are concepts resulting from the overall purpose of the cooperation proposals, which is mostly consistent with the motivation of Statistics Estonia. This means that, in our case, scientists are aware of the need to consider the perspective of the register-based census, which has been stimulating and enriching the cooperation. On the other hand, there is a professional interest in the census caused by traditions and other factors.

Problems that may hinder scientific collaboration are common to all institutions and related to the scarcity of resources (unable to participate or be involved), but more inherent barriers to cooperation are different and/or uncompromising approaches, e.g. the understanding of demographists and mathematicians-statisticians of data accuracy or methods of data cleansing.

The best results for the census project have been yielded by cooperation projects with research institutions that were focused on innovative solutions (index-based estimation of current population size and migration, also methodology for forming families from partners having wrong information on living-places). These solutions are useable also in other countries having similar problems, if there exists a good system of registers.

It is true that our cooperation with the research community has been fully based on the commitment to the census. Problems have been solved by going deep into the problems.

Our recommendation is to share information and clearly formulate the expectations of the cooperation partners, which would help to avoid communication problems and to be a reliable partner. It is necessary to encourage cooperation as a process that can be streamlined, but this requires agreeing on common goals at the very beginning of the cooperation. It is always necessary to understand the partners' motives, needs and to comprehend them, but the producers of statistics should be able to remain independent in their choices.

## 5. References

1. Puur, A., Sakkeus, L., Aben, S., (2013), Registripõhise rahva ja eluruumide loenduse (REGREL) metoodika väljatöötamine. Projekti lõpparuanne.Tallinn. In Estonian.
2. Lehto, K. (2014), REGREL-i pilootloendus.Tallinn. In Estonian.
3. Laast-Laas, J. (1996), Registrid ja klassifikaatorid rahvaloenduses. Automaatkodeerimine, AS Infoekspert.In Estonian.
4. Tiit,.E.-M., Maasing, E., (2016). Residency index and its application in census and population statistics. Quarterly Bulletin of Statistics Estonia.3/16. pp 53-60.
5. Tiit, E.- M., Vähi, M., (2017). Indexes in demographic statistics: a methodology using nonstandard information for solving critical problems. Papers on Anthropology XXVI/1 pp.72-87.
6. Statstics Estonia, (2016).Pressrelease no 48.
7. Memorandum, (2018).Seisukoht REL2020 ettevalmistamise kohta.Not published.In Estonian.
8. T i i t , E.-M, K. Me r e s , M. V ä h i , (2012). Rahvaloenduse üldkogumi hindamine. *Eesti Statistika Kvartalikiri*, nr 3, lk 79–95.
9. Äär, Helerin, (2017). Registripõhise ja tegeliku elukoha kattuvus. *Eesti Statistika Kvartalikiri*, nr 1, lk 73—83.
10. Tiit, Ene-Margit, Maasing, Ethel, Visk, Helle, Vähi, Mare, (2018).Rahvastikustatistika probleemkohad ja nende lahendamine registrite abil. RiTo Nr 38, lk 165-180.
11. Kütt, Kairiin ,(2014).Leibkonnad ja perekonnad registripõhises rahva ja eluruumide loenduses. Magistritöö. <http://dspace.utlib.ee/dspace/handle/10062/42537?show=full>
12. Visk, Helle, (2018). An index-based approach to determine partnership in a register-based census. IAOS Prize for Young Statisticians.
13. Tiit, Ene-Margit,(2017). Residency Testing. Estimating the true population size of Estonia. Statistics in Transition, Warsawa, 2, 211−226.
14. Maasing, Ethel; Tiit, Ene-Margit; Vähi, Mare,(2017). Residency index - a tool for measuring the population size. Acta et Commentationes Universitatis Tartuensis de Mathematica, 21 (1), 129−139.

## Maasing, Ethel (2015). Eesti alaliste elanike määratlemine registripõhises loenduses,http://hdl.handle.net/10062/47557

1. 2011. Aasta rahva ja eluruumide loendus. Metoodika. Tallinn, 2014.