

Problematic issues in preparation of register-based population census and ways of solving them

Ene-Margit Tiit, Statistics Estonia



Content

- Registers and registers' system
- Assessed population size and under-coverage of census
- Residency index
- Using residency index for defining foreign migration
- Family and partnership

Development of society and statistics

- The development of society brings changes in many phenomena and concepts.
- This makes producing and using statistics complicated and often ambiguous.
- Also preparing census using traditional concepts will be in some sense questionable.
- We will discuss several problems and show how Statistics Estonia has solved them using an innovative method that relies on cross-usage of a great number of registers, see [1].

REGISTERS AND REGISTERS' SYSTEM

Registers and the system of registers

Register is a data-set satisfying the following conditions:

- Each object belonging to register has a unique identifier
- The register is complete, i.e. it contains all data of the population (of this register).
- The information in the register is checked and updated regularly (at least once a year).

System of registers consists of a set of different registers, where

- All objects have the same identifiers in different registers;
- Different registers are linked by objects' identifiers.

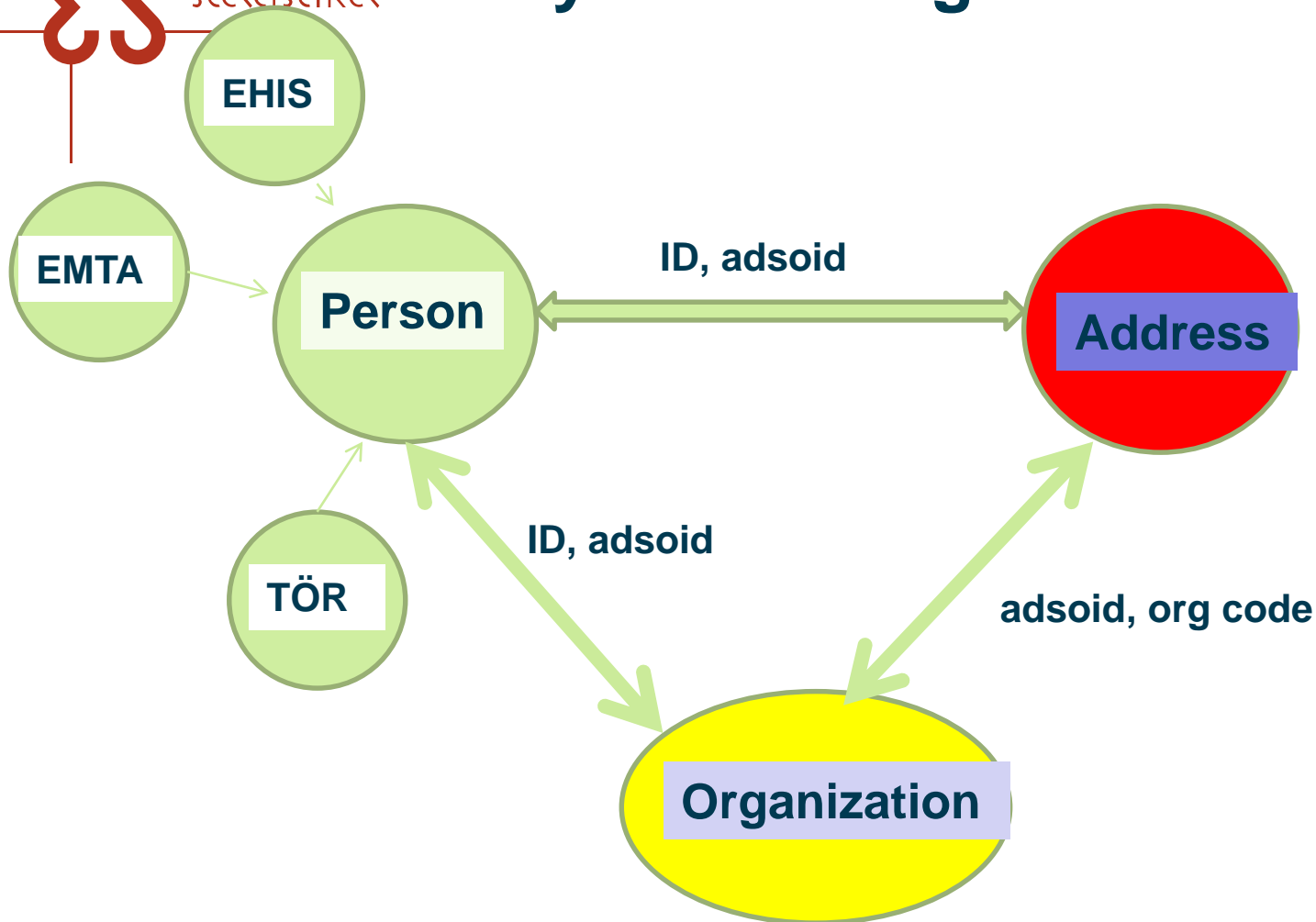


Register's system in Estonia

- Estonia has a system of registers that meets the above-mentioned requirements, and, in addition to persons and dwellings, it also contains businesses as objects with their identifiers.
- A precondition for register-based statistics is continuous checking of the quality of the data contained in the registers, particularly their correctness (correspondence between the data and the actual situation).
- A relatively efficient way for this is the cross-usage of registers, i.e. the data of one register are checked by means of other registers that contain independent data.



System of registers in Estonia



There are three central registers: Population register (persons), Buildings' registers (containing detailed addresses of dwellings and enterprises) and organizations and enterprises' register. Central registers are linked with sub-registers (education, work) and are also linked with each other using ID-codes for persons etc.

Finding the population size

The population size is usually found by means of a census. In recent times, however, under-coverage of the census has become a problem in many countries. This has several causes.

- First, people are much more mobile than earlier.
- The other reason, however, is psychological – people do not want to share their personal data.

This is a serious challenge for providers of official statistics.

There are several methods to check the coverage of the census.

- Efficient methods for assessment of the size of the population can be used in countries with functioning system of high-quality registers, see [2—3].

ASSESSED POPULATION SIZE AND UNDER-COVERAGE OF CENSUS

Correction of census under-coverage in Estonia

- The latest Estonian census until now, in 2011, was under-covered.
- Signals about this came from census participants themselves; the number of people counted was almost 5% smaller than the number of people residing in Estonia according to the population register (PR).
- To get the actual number of population and a list of people actually residing in Estonia, a decision was made to use several registers containing personal data.
- This was actually the first attempt to correct population data by cross-usage of register data, see [5, 6].

- Data on the inhabitants of Estonia were collected from more than ten registers and sub-registers.
- Considering that people's activity in registers depends greatly on their age and gender, twelve separate gender-age groups were formed.
- In each group a logistical regression model was created to assess the probability that the person belongs to permanent population of Estonia.
- The model was applied to all persons registered in Population Register, but not enumerated.
- In conclusion, it turned out that the 2011 census was under-covered 2.3% [6]. From the year after the census, population statistics used corrected census data.
- It was also revealed that the number of Estonia's residents in the population register was overestimated by almost 3%, the reason for it being under-registration of departures from Estonia.

RESIDENCY INDEX



Finding the population number by means of residency index

- The task was – keep the list of permanent residents and the list of persons who have left the country. Both sets together form the extended population; these are the persons about whom there are (have been) data in Estonian registers.
- To define permanent residents, the residency index $R(j,k)$ was created according to the formula

$$R(j, k) = d * R(j, k - 1) + g * X(j, k - 1), \quad (1)$$

where (1) is the value of the residency index of person i in year k , ($0 \leq R(j,k) \leq 1$), see [3, 4].

- In formula (1) $R(j, k - 1)$ is the residency index in the previous year, $X(j, k - 1)$ is the weighted sum of signs of life collected by person i in the previous year,

$$X(j, k - 1) = \min(1, \sum a_j b_j). \quad (2)$$

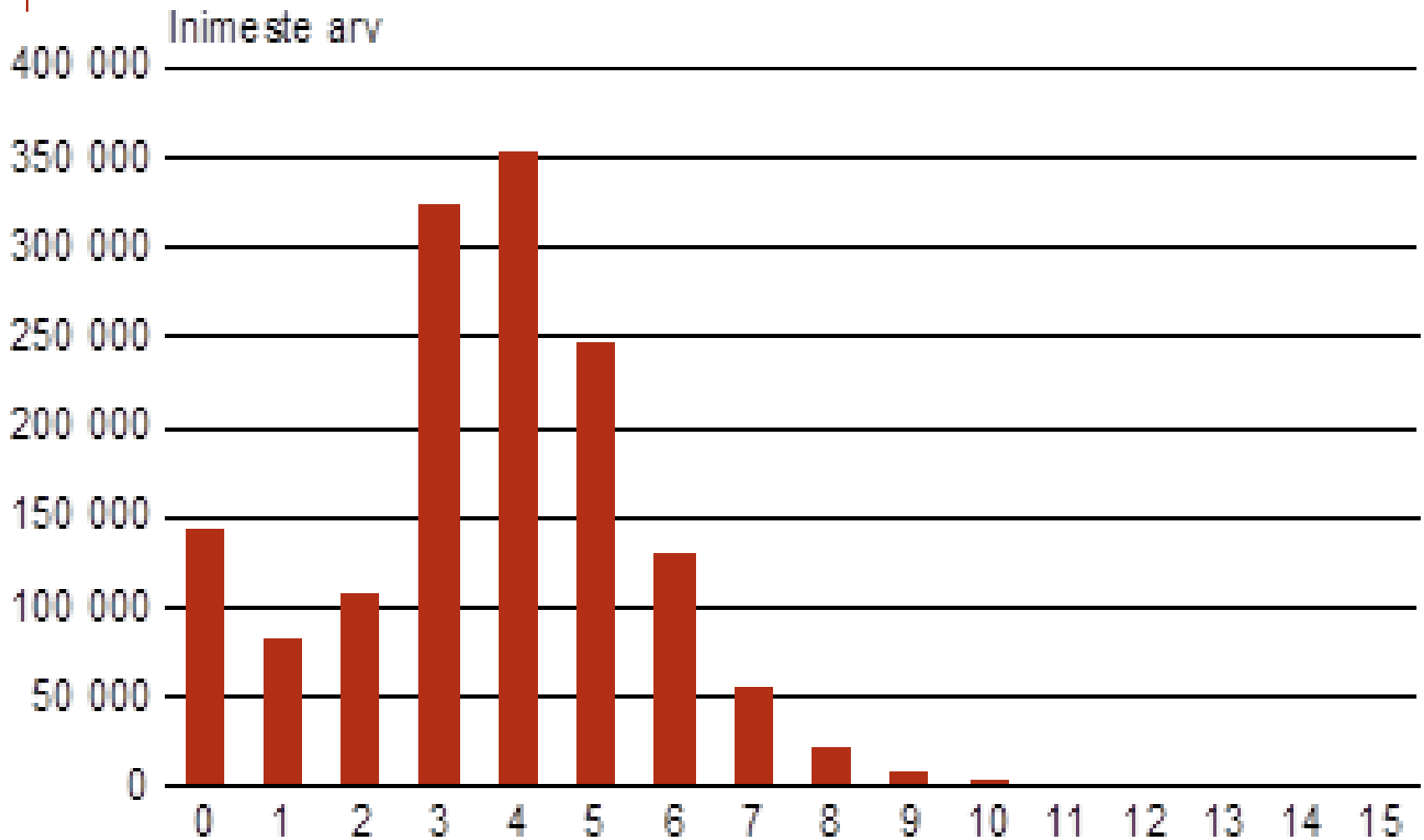
- Sign of life b_j corresponding to register j is a binary variable the value of which is 1 if the person i has been active in register and 0 if the person has not been active in this register.

- So, the person acquires a corresponding sign of life if s/he, for example, visits a family physician, studies at a school in Estonia, receives social security benefits from the local government or buys a prescription drug etc.
- Number of possible registers and sub-registers in Estonia is more than 20.
- The weights a_j of life signs b_j can be calculated in different ways, for instance, using some statistical procedures (e.g. logistic regression) and test-samples.
- The parameters d , g in the formula (1) and the threshold c can be calculated using empirical data.
- The residency index is calculated each year for all the persons belonging to the extended population, and, in year k , persons whose residency index is higher than threshold c are considered to be residents.



eesi
STATISTIKA

Distribution of values of residency index in the case of Estonian population



USING RESIDENCY INDEX FOR DEFINING FOREIGN MIGRATION

Assessment of migration

- One of the most complicated task is defining foreign migration Traditionally, information on it is received from officially registered migration statistics, which tends to underestimate the number of migrants.
- Many emigrants do not register leaving the country for several reasons – forgetting or they wish to keep the status of a permanent resident in order to use its benefits.
- If foreign migration is not balanced, the error cumulates during years and distorts the total number of emigrants and also the number of population.
- Such was the situation in Estonia before the 2011 census when, because of under-registered emigration, the number of population in PR was overestimated by almost 3%, see [1, 4].

Using the residency index the migration can be defined in the following way:

- Immigrants – people, who were not residents in year $k-1$, but are residents in year k .
- Emigrants – people, who were residents in year $k-1$, but are residents in year k .
- There are some additional details (e.g. birth and death processes) to make the rule more exact and suitable for calculations.
- Index-based foreign migration includes both registered and unregistered migration and therefore, index-based migration has greater volume than traditionally calculated migration.
- Using the residency indexes it is also possible to define **pendulum migration** and **transnationality**.

FAMILY AND PARTNERSHIP



Household and family

- Household is in register-based statistics the set of people living in the same dwelling. The relations between these people do not have any sense.
- Family is more significant concept, it has three forms:
 - Married or cohabiting couple;
 - Married or cohabiting couple with child or children;
 - Single parent with child or children.

If the couple is married or their cohabitation is legally registered, then it is easy to recognise the family. But in Estonia about 30% of couples are not registered cohabiting. To recognise them there are two ways.

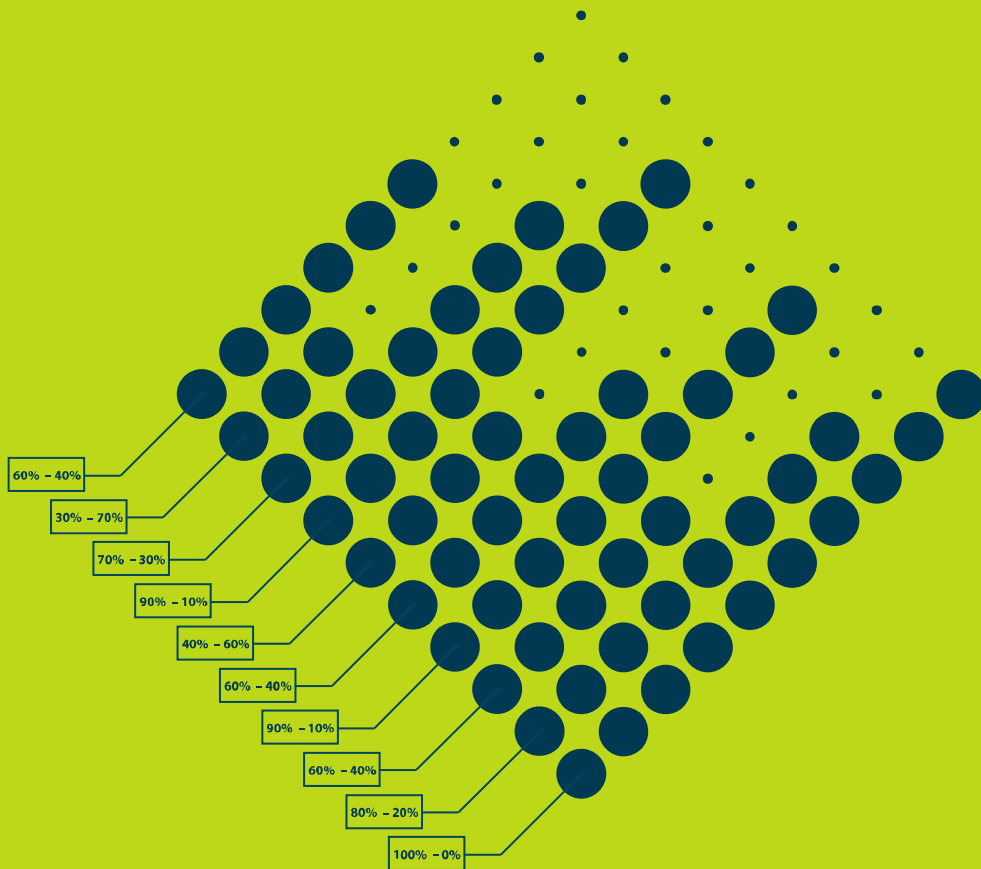
- If this couple is registered in the same dwelling then it is quite easy to determine the family, its type and fix for all family members their status in household and family, see [7].
- Otherwise, there exists register-based methodology for defining the couples using partnership index, see [8].

Summary

- The development of official statistics is densely connected with using data from alternative sources – registers and also big data, that time by time substitute traditional statistical methods of data collection.
- At the same time it is necessary to take into consideration that using new data sources needs also new data analyzing and decision making methodologies, see [9].



1. Tiit, Ene-Margit, Maasing, Ethel, Visk, Helle, Vähi, Mare (2019) Rahvastikustatistika probleemkohad ja nende lahendamine registrite abil. RiTo Nr 38, 165--180
2. Tiit, Ene-Margit (2017). Residency Testing. Estimating the true population size of Estonia. *Statistics in Transition*, Warsawa, 2, 211–226.
3. 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
4. Tiit, Ene-Margit; Maasing, Ethel (2016). Residency index and its applications in censuses and population statistics. *Eesti Statistika Kvartalikirj*, 3, 41–60.
5. Tiit, E.-M., Meres, K., Vähi, M. (2012). Rahvaloenduse üldkogumi hindamine. *Eesti Statistika Kvartalikirj*, nr 3, lk 79–95
6. Tiit, Ene-Margit (2012). 2011. aasta rahva ja eluruumide loenduse alakaetuse hinnang. *Eesti Statistika kvartalikirj* nr.4, lk 110—115
7. Kütt, Kairiin (2014) Leibkonnad ja perekonnad registripõhises rahva ja eluruumide loenduses. <http://dspace.utlib.ee/dspace/handle/10062/42537?show=full>
8. Visk, Helle (2019) An index-based approach to determine partnership in a register-based census. *Statistical Journal of the IAOS*, vol. 35, no. 2, pp. 245-251, 2019.
9. Tiit, E.-M.; Vähi, M. (2017). Indexes in demographic statistics: a methodology using nonstandard information for solving critical problems. *Papers on Anthropology*, 26 (1), 72–87



Thank you for attention!