

Residency testing

Ene-Margit Tiit

Residency problem. Case of Estonia

Number of residents or population size is important for all countries and regions, especially nowadays, when population is mobile and results of censuses expire rapidly.

From the time when different registers were created and implemented, the situation has changed, as the number of residents can be counted also from registers. In countries having population register and/or some other good (administrative) registers, the population size can be calculated at any time without interviewing the people. Still, the situation is not so simple as the registers can be over- and under-covered. For instance, in Estonia after population and household census 2011 (PHC2011) we had three different numbers of population size differing from each other for several percents:

- Size of census population – 1 294 455.
- Population size calculated using registered population events and population size of census PHC2000 – 1320 000.
- Number of Estonian residents in Estonian Population Register (EPR) – 1365 000.

In some age-groups the difference between various estimates was almost 10%. It became evident that census population was somewhat under-covered – what is very common nowadays, when the people are mobile and the migration between countries belonging into EU and/or Schengen group is free. Also, it seemed that the population size fixed in EPR was somewhat over-covered. In 2012, immediately after PHC2011, the true size of Estonian population was estimated by loglinear models using register data as explanatory variables. About 30 000 persons (2.3% of population) were added to census population to get the official population for demographic calculations. Each added person was identified by his/her ID-code.

Preparation for PHC2020. Estimation of census population

As Estonia has a quite well-functioning registers' system, it has been decided that the following population and household census in 2020/2021 will be organised without personal enumeration and interviewing, but based on using registers, as this has been made already in several countries. That means, it is necessary to know beforehand the census population – the identified set of residents. The task will be more complicated, when it will be necessary to estimate the number of residents each year – as it will be in future, assuming the regulations of Eurostat.

To solve this problem, a concept of residency index was elaborated. Residency index $R(i, k)$ is a characteristic, calculated for each year k for each person i from the set of potential residents and having the value between 0 and 1. For certain residents the index has value 1, for certain non-residents – value 0. All persons having the index higher than threshold c are considered as residents in year k . Hence, the index can be regarded as probability of being a resident.

In the beginning of each year the residency index is recalculated using the following formula:

$$R(i, k + 1) = dR(i, k) + gX(i, k), \quad (1)$$

where d and g are fixed constants ($0 \leq d, g \leq 1$), $R(i, k)$ is the person's index in previous year and $X(i, k)$ is the sum of *signs of life* of the person i in year k , defined by the following formula:

$$X(i, k) = \sum_{j=1}^J x(i, k, j), \quad (2)$$

where the binary variable $x(i, k, j)$ is so-called *sign of life* of the person i in the register j , having the value 1, if the person has been active in the register j at least once during the year k and 0 else. The number of registers available is J . In case of Estonia J had value 24 that means, we used for calculations 24 administrative registers and sub-registers. To assure that index does not have its value higher than 1, the value of index is truncated to 1 after using the formula (1).

The initial values of parameters c (threshold), d (stability parameter and g (signs of life parameter) were calculated theoretically, assuming the appropriate time for changing the residency status, that means, when a resident loses his/her residency because of lack of signs of life and when a person (not formally immigrated) will be considered as a resident because of having lots of signs of life. The estimated values were checked statistically.

The initial year for all calculations was 2011, the year of census and the initial values of $R(i, 0)$ were defined using census data.

The population sizes for years 2012—2014 were estimated using the traditional population statistics (by population events) and also with help of residency index. The difference between two estimates was less than 1%, where the size calculated by index was somewhat larger than the size calculated by traditional methodology.