Overview

- Introduction
  - Purpose
  - Available data
- Logistic regression
- Comparison
  - Linear regression
  - Logistic regression
  - Discriminantory analysis
  - Published population
- Conclusion and plans for future
Introduction

- **Population and household census (PHC)** aim is to collect data on the entire national population, households and dwellings on a fixed point in time.

- The next PHC in Estonia is intended to be fully register-based (end of 2020).

- The aim of this work is to determine Estonian population using only administrative registers at the end of 2014.
  - This work does not deal with the full addresses, only on a country level.
Population Register (PR)

- PR contains data about
  - all Estonian citizens;
  - foreigners, who have registered their address in Estonia;
  - foreigners, who have got an Estonian residence permit.

- Everybody is obliged to register their right address to PR by the law.

- **BUT:**
  - people who have left Estonia do not register their leaving in PR;
  - people who have come (back) into Estonia do not give this information to PR.
Important points

- It can be assumed that the people, who actually live in Estonia are represented in other administrative registers, because they are using services and receive payments.

- All persons have a unique identification code that is also used by all administrative registers.
  - ID codes are replaced with anonymous codes that are used only by SE.
Administrative registers

- Estonian Education Information System;
- Register of Social Services and Benefits;
- Health Insurance Database;
- National Defence Obligation Register;
- State Pension Insurance Register;
- Register of persons registered as unemployed or job-seekers, and of provision of labour market services;
- Register of Residence and Work Permits;
- E-file system (crime documents, court documents etc);
- Estonian Traffic Register (Changes of driver’s licenses, changes of vehicles);
- Register of Employment.
ASSUMPTIONS AND RESULTS OF LOGISTIC REGRESSION
Choosing the control groups

- **Definite resident**
  - In PR registered address is in Estonia and they were counted in the last census as a resident;
  - Born at 2012-2014;
  - Migrated to Estonia in 2012-2014.

- **Definite nonresident**
  - In PR registered address is not Estonia and they were counted as left for abroad in the last census;
  - In PR registered address is not Estonia and they were not counted in the last census.
Choosing age and/or sex groups

The average number of occurrences in registers

- Black: Age groups limits
- Blue: Males
- Red: Females

24-28 August 2015

Ethel Maasing
## Results of logistic regression

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>0 - 6</th>
<th>7 - 16</th>
<th>17 - 22</th>
<th>23 - 30 males</th>
<th>23 - 30 females</th>
<th>31 - 45 males</th>
<th>31 - 45 females</th>
<th>46 - 62 males</th>
<th>46 - 62 females</th>
<th>over 63</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Count of reg</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Residence permit</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Education</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Social benefits from local mun</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Servicemans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver license</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Changing vehicle</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>E-file</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Special welfare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Social benefits from the state</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Pensions</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Family support</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Parental benefit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dental support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital recipe</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment figures</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Incapacity of work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Health insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
**Sensitivity and specificity**

- **Sensitivity** shows the proportion of residents that the used model predicts correctly;
- **Specificity** shows the proportion of nonresidents that the used model predicts correctly.

<table>
<thead>
<tr>
<th></th>
<th>0 - 6</th>
<th>7 - 16</th>
<th>17 - 22</th>
<th>23 - 30 males</th>
<th>23 - 30 females</th>
<th>31 - 45 males</th>
<th>31 - 45 females</th>
<th>46 - 62 males</th>
<th>46 - 62 females</th>
<th>over 63</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>98,2%</td>
<td>98,4%</td>
<td>96,4%</td>
<td>92,2%</td>
<td>94,9%</td>
<td>93,1%</td>
<td>98,0%</td>
<td>95,9%</td>
<td>98,2%</td>
<td>99,4%</td>
</tr>
<tr>
<td>Specificity</td>
<td>97,1%</td>
<td>99,6%</td>
<td>98,2%</td>
<td>96,3%</td>
<td>98,2%</td>
<td>97,3%</td>
<td>99,0%</td>
<td>97,3%</td>
<td>99,0%</td>
<td>98,6%</td>
</tr>
</tbody>
</table>

- Most difficult with 23–45 years old males;
- Simplest with 7–16 years old.
COMPARISON WITH OTHER METHODS AND PUBLISHED POPULATION
Information about data

- Different assumptions for methods
  - 6 age and/or sex groups
  - Immigrants in period 2012-2014 are not definite residents

- Published population is person-based
  - Results of the 2011 census
  - + undercoverage
  - + registered events (births, deaths and migration)
# Comparision

<table>
<thead>
<tr>
<th>Sum of res</th>
<th>Published population</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nres</td>
<td>res</td>
</tr>
<tr>
<td>0</td>
<td>125 869</td>
<td>32 622</td>
</tr>
<tr>
<td>1</td>
<td>5 714</td>
<td>23 768</td>
</tr>
<tr>
<td>2</td>
<td>5 944</td>
<td>7 477</td>
</tr>
<tr>
<td>3</td>
<td>2 364</td>
<td>8 754</td>
</tr>
<tr>
<td>4</td>
<td>13 089</td>
<td>1 237 258</td>
</tr>
<tr>
<td>Total</td>
<td>152 980</td>
<td>1 309 879</td>
</tr>
</tbody>
</table>
Published population = resident
Sum of res ≤ 3

24-28 August 2015
Ethel Maasing
Published population = resident
Sum of res ≤ 3
31-45 males
Conclusion and plans for future

- All registres were statistically important

- To add more registres
  - Register of Identity Documents
  - Register of Prisoners

- This method cannot be used later on because the control groups were based on the last census.

- Residence index – Ene-Margit Tiit
Thank you!