

First Results in Determining Permanent Residency Status in Register-Based Census

Ethel Maasing
Statistics Estonia

4th Baltic-Nordic Conference on Survey Statistics, 24-28 August 2015
Helsinki, Finland





Overview

- Introduction
 - Purpose
 - Available data
- Logistic regression
- Comparison
 - Linear regression
 - Logistic regression
 - Discriminatory 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

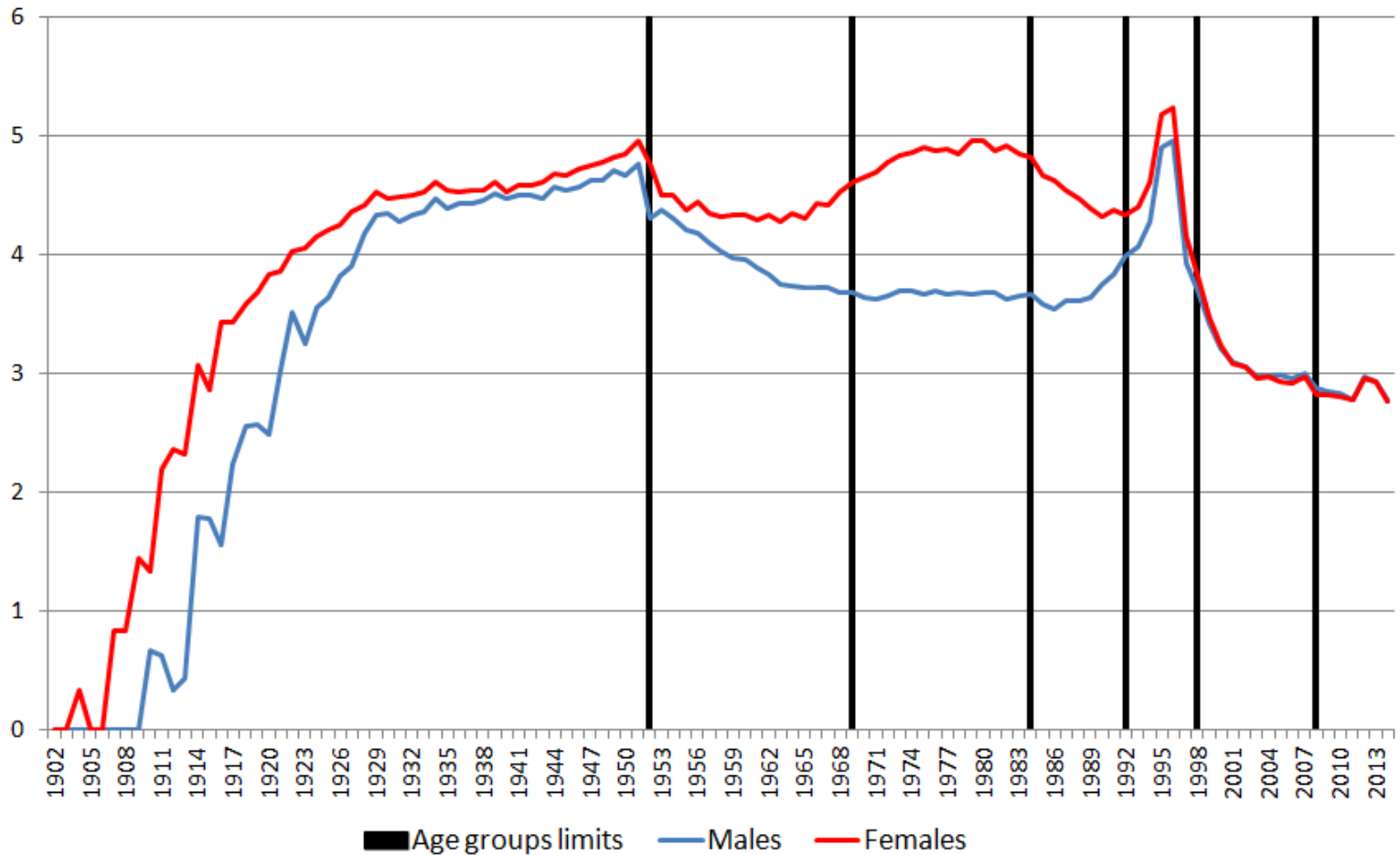
- Definite resident
 - In PR registered address is in Estonia and they were counted in the last census as a resident;
 - Born at 2012-2014;
 - Migrated to Estonia in 2012-2014.

- 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





Results of logistic regression

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



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.

	0 - 6	7 - 16	17 - 22	23 - 30 males	23 - 30 females	31 - 45 males	31 - 45 females	46 - 62 males	46 - 62 females	over 63
Sensitivity	98,2%	98,4%	96,4%	92,2%	94,9%	93,1%	98,0%	95,9%	98,2%	99,4%
Spesificity	97,1%	99,6%	98,2%	96,3%	98,2%	97,3%	99,0%	97,3%	99,0%	98,6%

- 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)

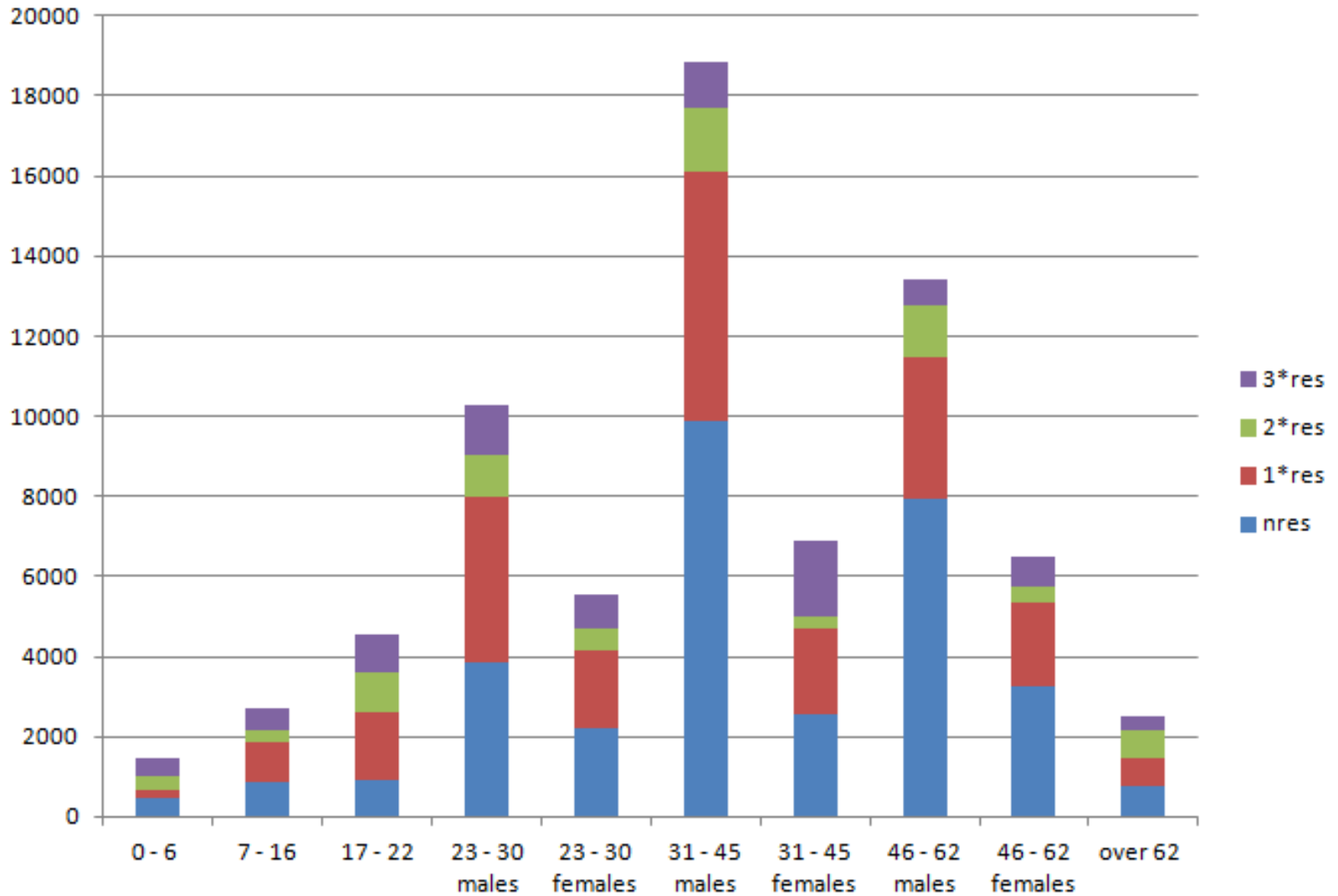


Comparision

Sum of res	Published population				Total	
	nres		res			
0	125 869	8,6%	32 622	2,2%	158 491	10,8%
1	5 714	0,4%	23 768	1,6%	29 482	2,0%
2	5 944	0,4%	7 477	0,5%	13 421	0,9%
3	2 364	0,2%	8 754	0,6%	11 118	0,8%
4	13 089	0,9%	1 237 258	84,6%	1 250 347	85,5%
Total	152 980	10,5%	1 309 879	89,5%	1 462 859	100,0%



Published population = resident Sum of res ≤ 3

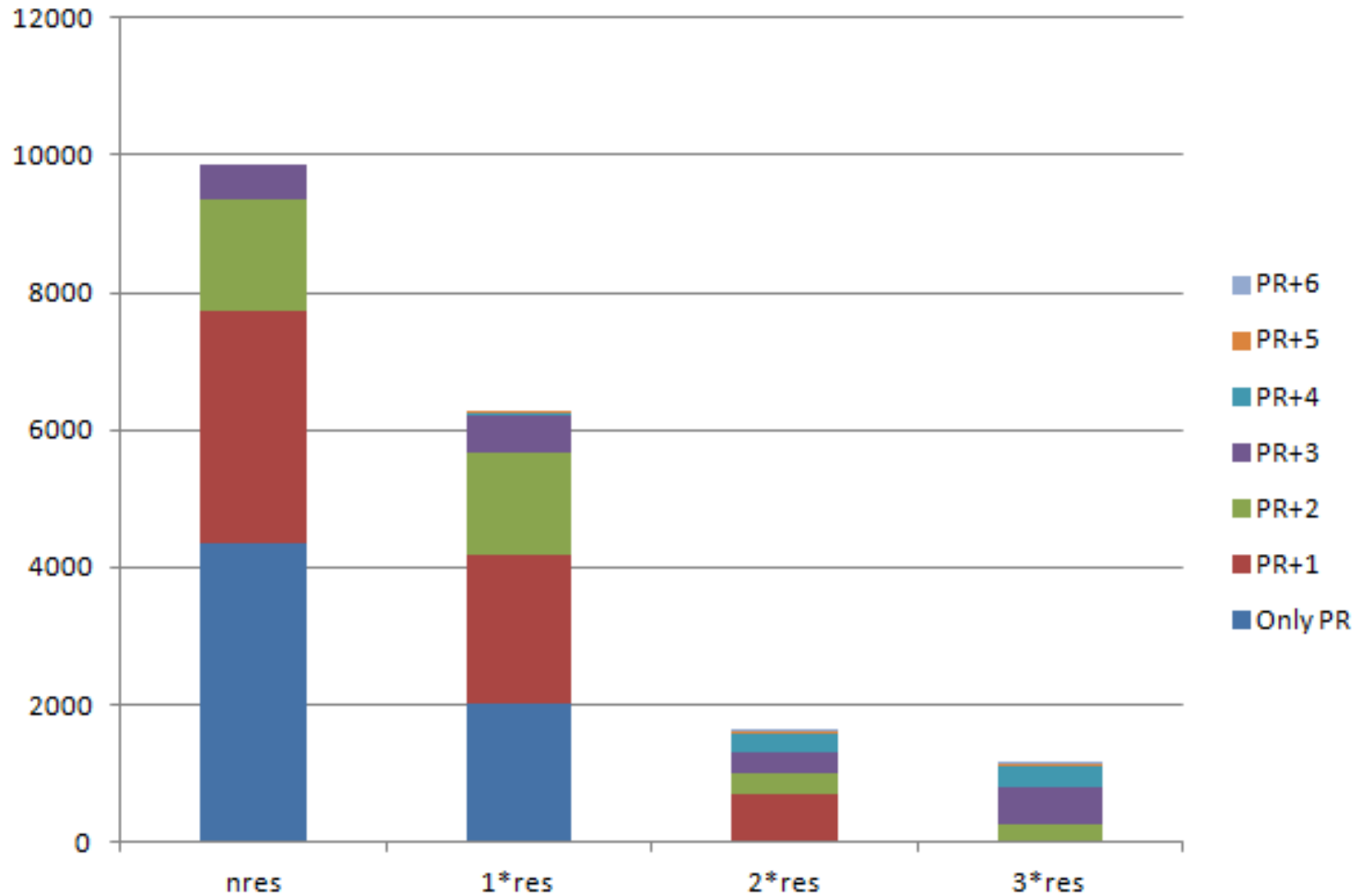




Published population = resident

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 can not be used later on because the control groups were based on the last census.

- Residence index – Ene-Margit Tiit

Thank you!

