A register-based input database for EUROMOD micro-simulation model

BaNoCoSS 2019
16-20 June

Kaja Sõstra

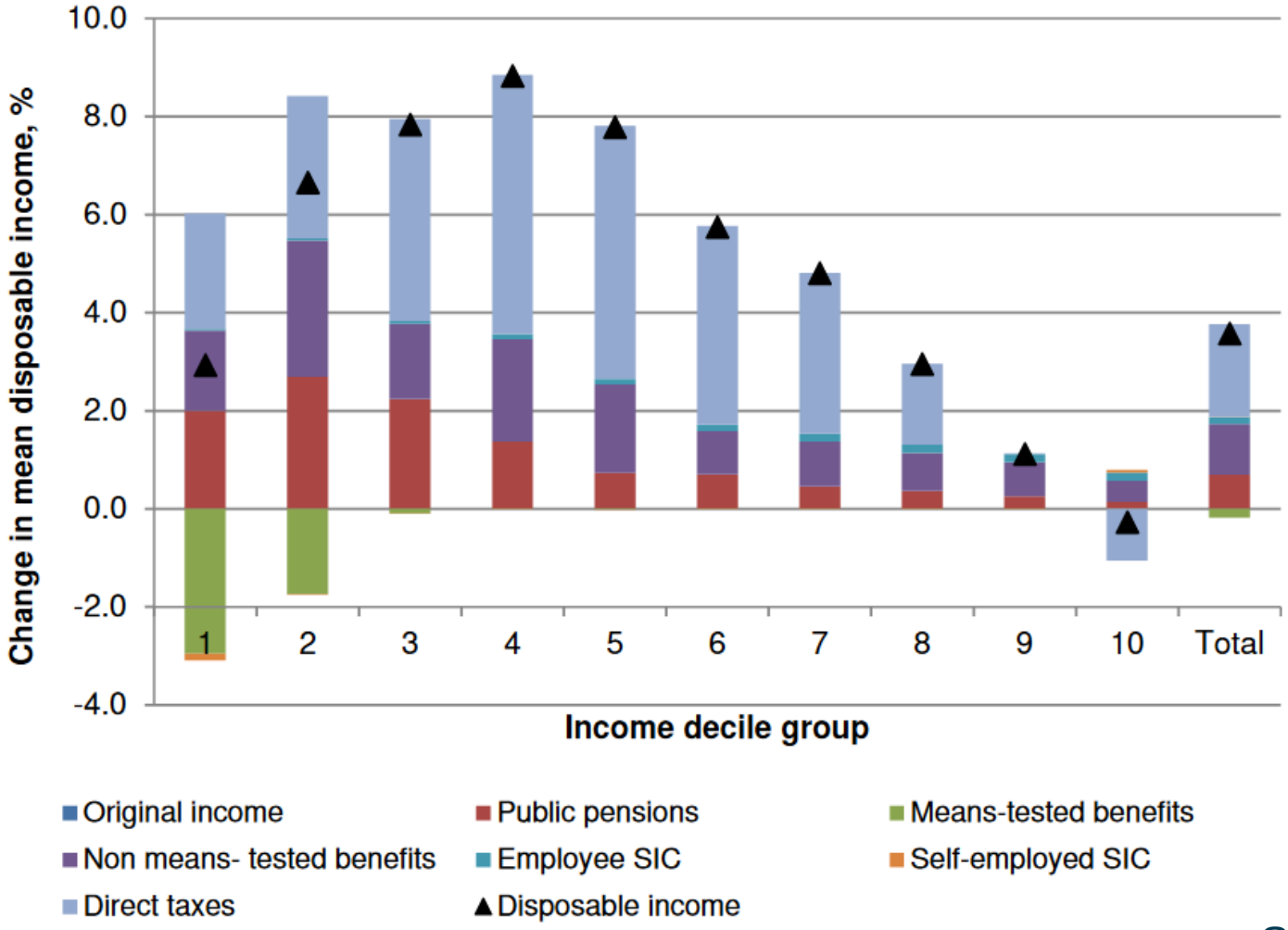
19.06.2019

eesti statistika

What is EUROMOD?

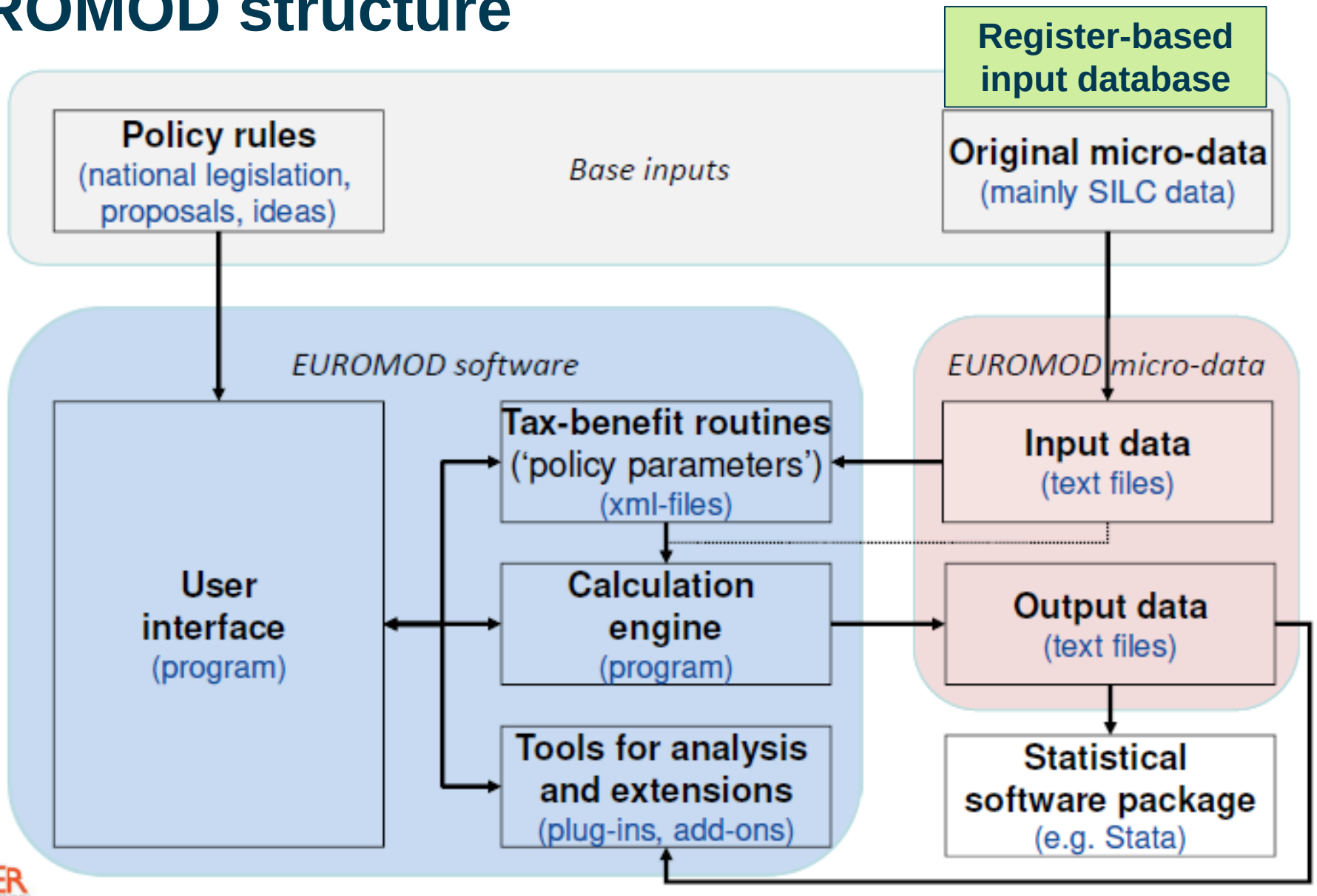
- Tax-benefit microsimulation model for the EU
- Enables researchers and policy analysts to calculate the effects of taxes and benefits on household incomes and work incentives for the population of each country and for the EU as a whole
- Developed by Institute for Social and Economic Research, University of Essex
- More information: https://www.euromod.ac.uk/

Example: Policy effects in 2017-2018, %



Source: Masso et al. (2018)

EUROMOD structure



Triggers for register-based input

- Strong interest to use EUROMOD model by Estonian analysts for analysing the impact of changes of taxes or benefiits
- Problems:
  - Accuracy – EU-SILC sample data is used for input
  - Scope and level of policy data – information about small groups is missing due to the sample
  - Time-lag for input data – 2016 data added in 2018
- Parliament elections in March 2019 – need to analyse the impact of party's promises

Prerequisites

- Experience and methodology for using register data:
  - Income variables in SILC are mainly based on the register data since 2014
  - Methodology developments for register-based census can be used for other register-based projects – residency index, partnership index, algorithms for calculating census variables
- Register data was available in SE, no new data requests from registers
- Human resources – one full-time working person
- Cooperation with EUROMOD national team

Steps of project

- Joint project with Center for Policy Studies PRAXIS, financed by Ministry of Economics
- Feasibility study – February-June 2018
- Constructing input micro-data based on registers covering income and taxis of whole population in 2017 – October 2018 – January 2019
- Validation and amending the micro-data – February – April 2019
- Use by specialists of Ministry of Finance – April 2019

Content of dataset

Topic	Number of variables	Examples of variables
socio-demographic	24	sex, age, place of residence, citizenship, marital status, disability
identifiers	6	
labour market	18	activity status, number of months spent in full/part-time work, economic activity, number of months in retirement
income	34	employee income, income from self-employment, number of month receiving employee income
benefits	34	education-related allowances, unemployment benefits, sickness benefits, disability benefits, old-age benefits
taxes	11	Personal income tax, property tax
expenses	5	mortgage payment : interests
assets	9	number of rooms available to the household, car ownership

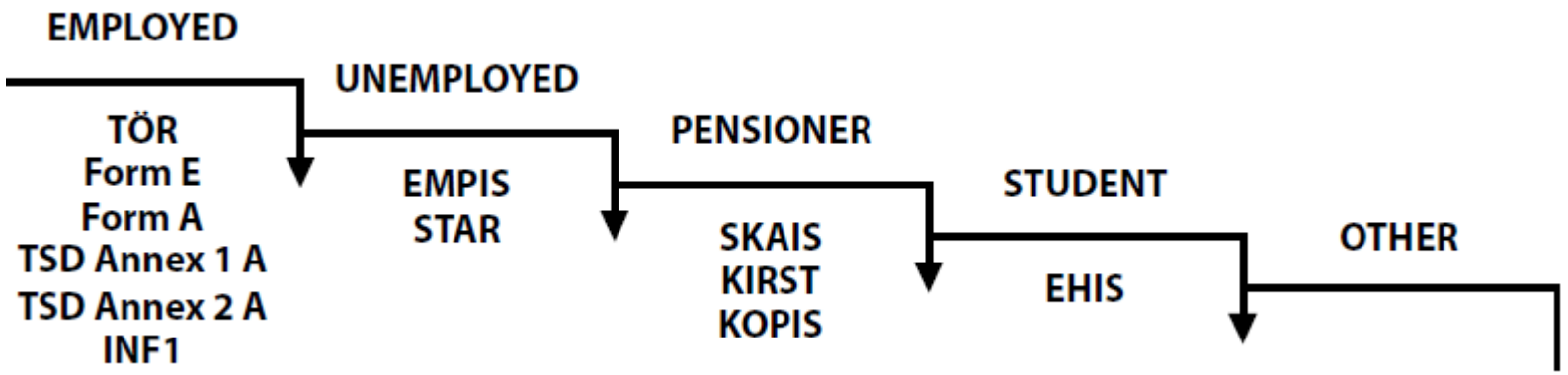
Registers used for constructing dataset

- Address Data System
- Commercial Register
- Estonian Education Information System
- State Register of Construction Works
- Estonian Unemployment Information System
- Health Insurance Information System
- Register of Mandatory Funded Pension
- Land Register
- National Defence Obligation Register
- Register of Taxable Persons
- Estonian National Pension Insurance Register
- Population Register
- Social Services and Benefits Registry
- Register of Employment
- Estonian Traffic Register

Constructing variables – activity status

- Questionnaire-based
  - Your main activity status – full/part-time employee, self-employed, unemployed, student, retired, disabled etc

- Register-based



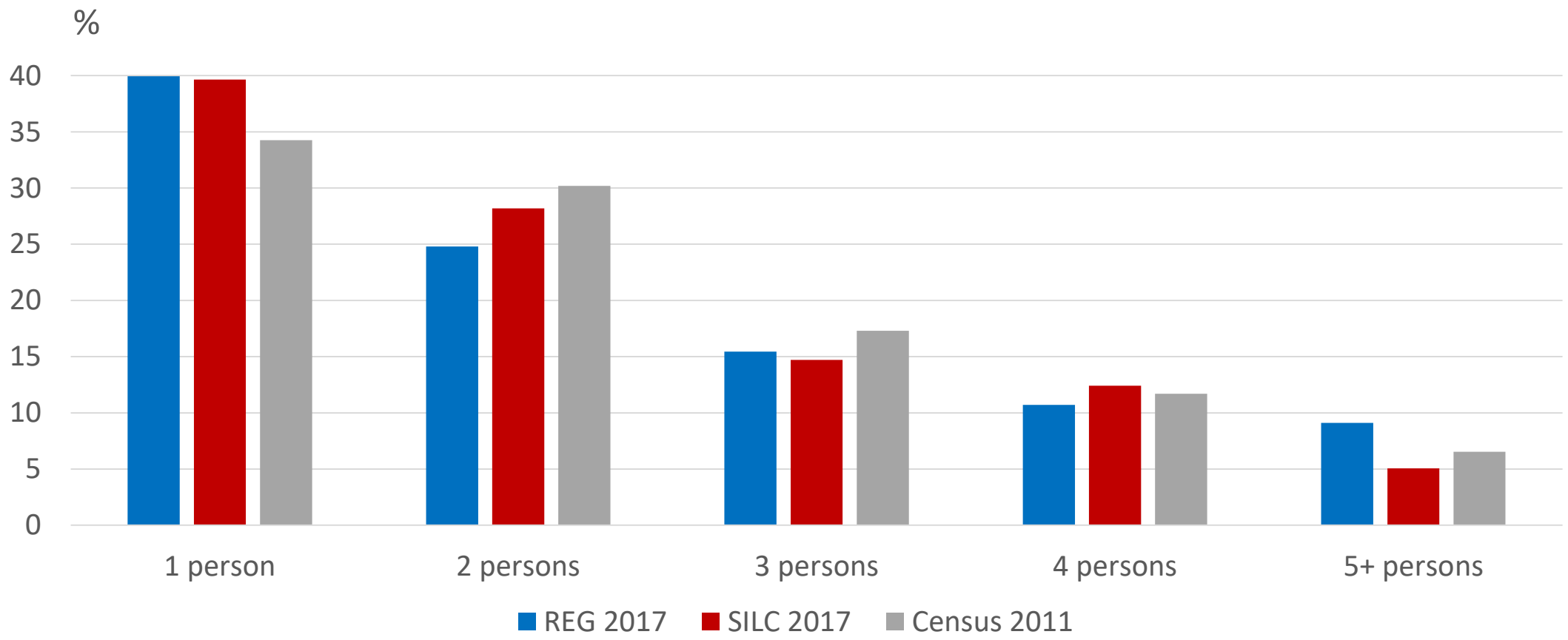
Problems

- Metadata of register data is missing or incomplete
- Missing data, erroneous values in register data
- Incomplete data from registers – missing units
- It is not possible to construct all variables based on register – use of company car, housing costs, municipality benefits
- Registered addresses differ from real address -> register-based households differ systematically from real situation

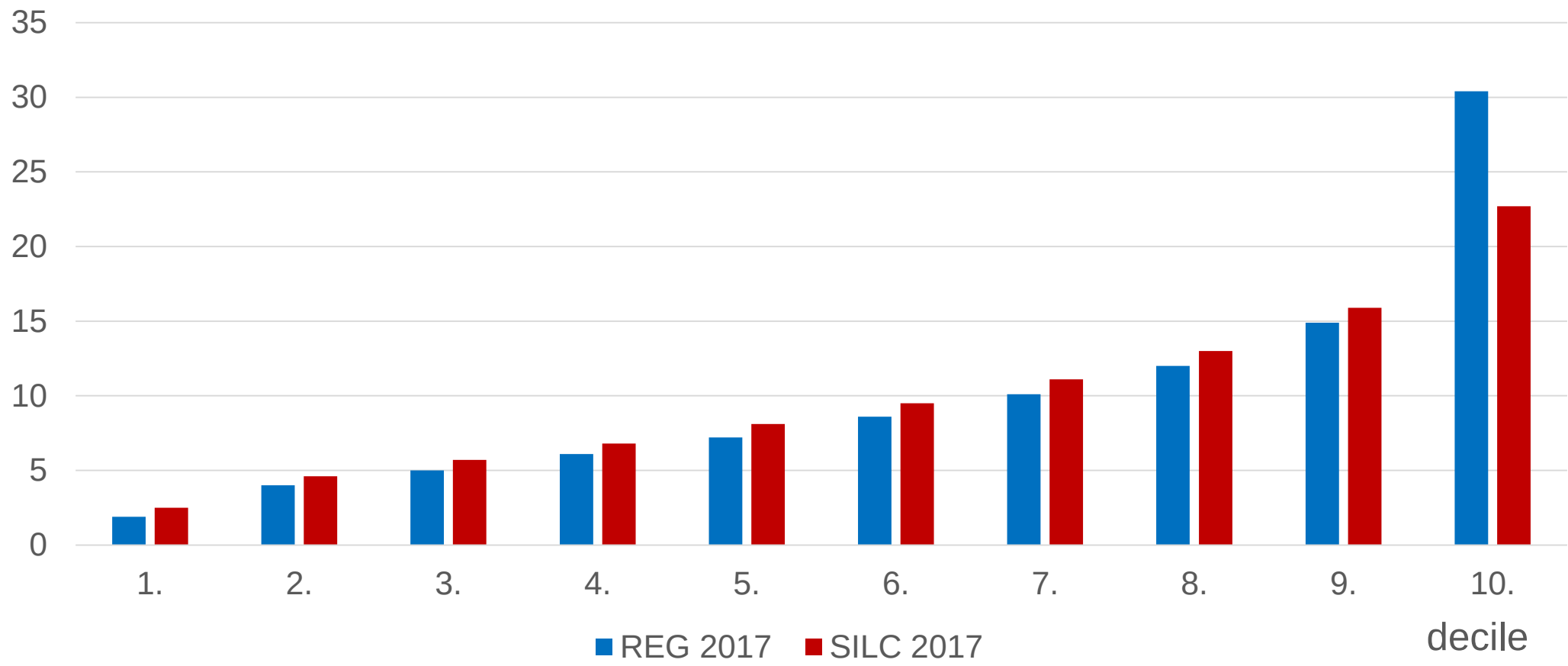
Validation of the dataset

- Dataset was macro-validated using different independent data sources:
  - SILC data for comparing the distribution of income
  - LFS for the number of employed persons
  - Statistics on benefits – the total sum and the number of recievers
  - Other relevant statistics

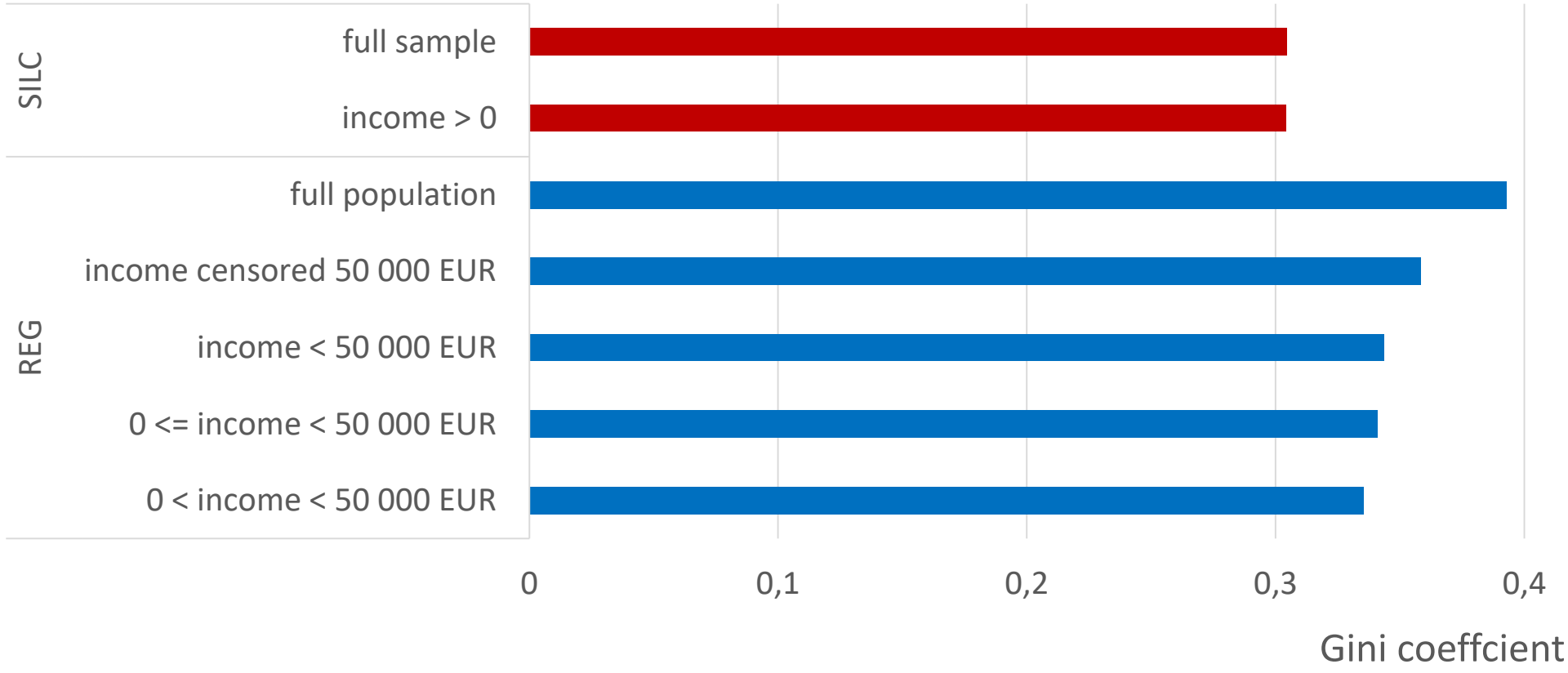
Size of household



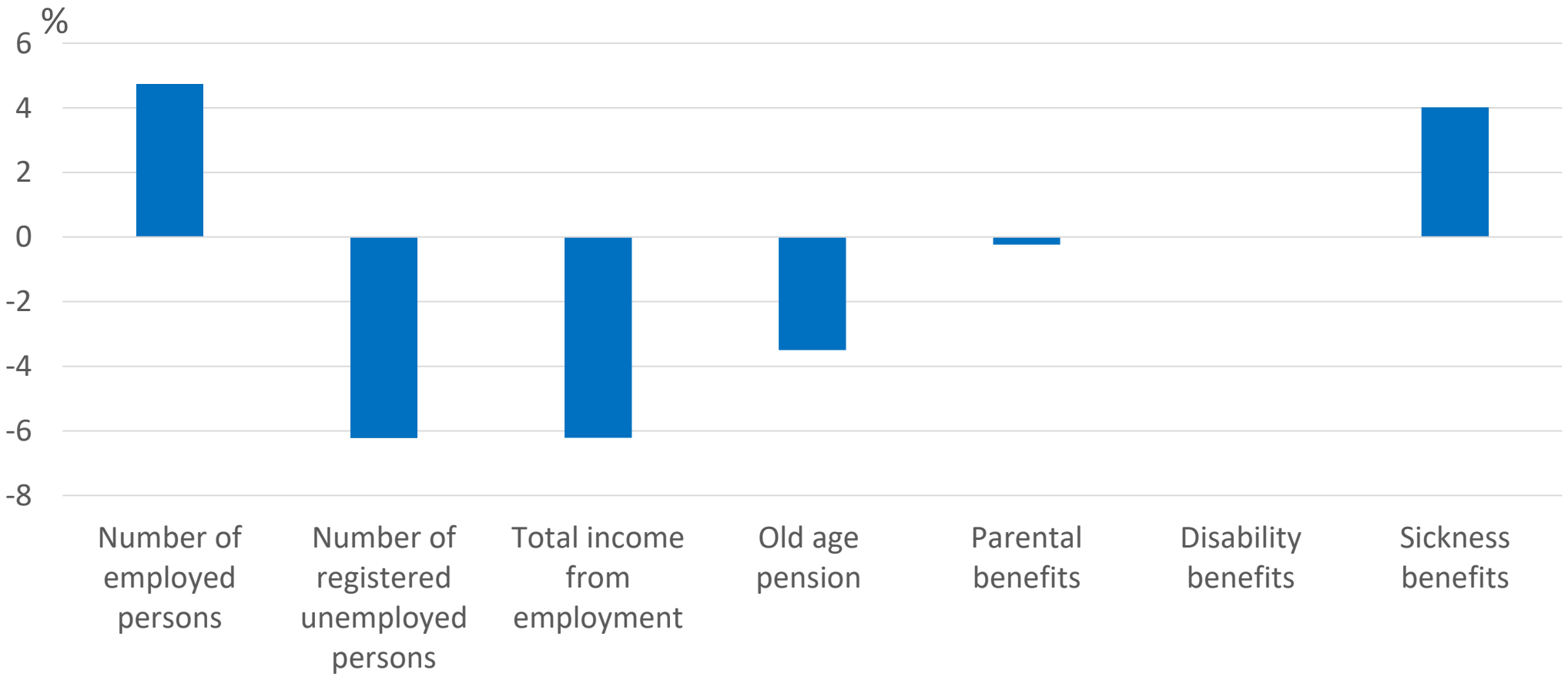
Distribution of disposable income by deciles, %



Impact of extreme values of income to Gini coefficient



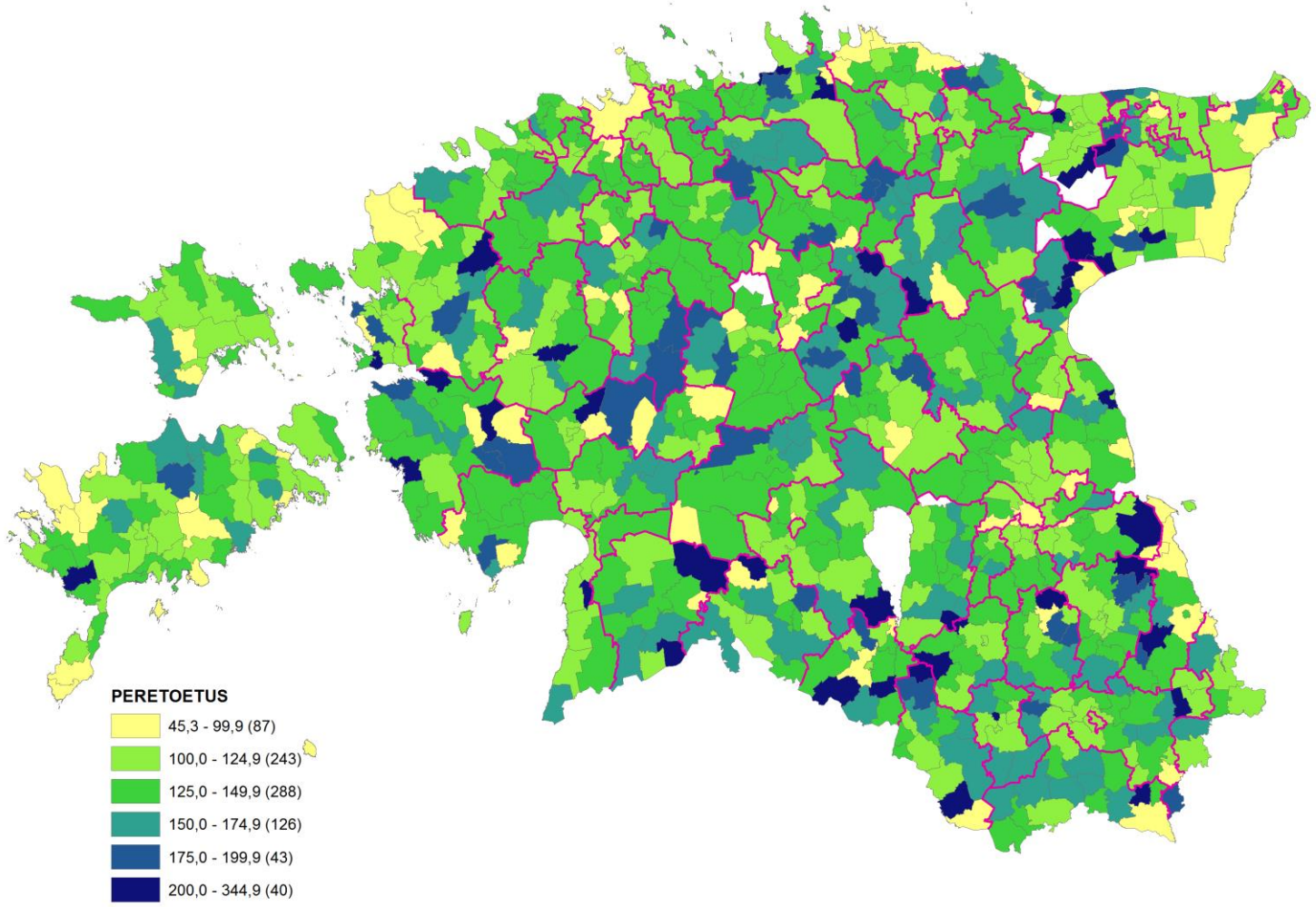
Difference between EUROMOD dataset and comparison statistics, %



EU-SILC vs register based input

	EU-SILC/Estonian-SILC	Administrative registers
Population coverage	Sample of ca 15 thousand persons	Whole population (1.3 mln)
Time lag	2-3 years	1-2 months
Panel data features	Partial (rotating panel)	Complete
Residency & household structure	Complete	Partial
Socio-demographics	Mainly current (at time of interview)	Annual or monthly
Incomes	Previous calendar year (mainly register-based), non-cash incomes	Annual or monthly (100% register-based), only cash incomes
Tax deductions	-	Detailed (annual)
Consumption	Partial (housing, child care)	-
Assets	Partial (housing, cars)	Detailed (domestic only)
Quality problems	Non-response, measurement errors	Incomplete or out-of-date information, income underreporting

Additional statistics for small regions
Family benefits by locality, euro



Conclusions and future plans

- The first trial for constructing the dataset of income of whole population was successful
- Dataset can be used for EUROMOD model and ohter purposes
- Next steps
  - to construct datasets for 2018 and 2013-2016
  - to develop monthly updates of dataset

References

Masso, M., Paulus, A., Piirits, M., Biin, H., Melesk, K. (2018). Estonia 2015-2018. EUROMOD Country Report. https://www.euromod.ac.uk/sites/default/files/country-reports/year9/Y9_CR_EE_Final.pdf

Paulus, A. (2018). „Feasibility study of constructing and implementing register-based input dataset for EUROMOD Estonian model“ (in Estonian). Tallinn: Center for Policy Studies PRAXIS.

Paulus, A., Piirits, M (2019) „Project report: EUROMOD model with Estonian register data“ (in Estonian). Tallinn: Center for Policy Studies PRAXIS

Sutherland, H. & Figari, F. (2013). „EUROMOD: the European Union tax-benefit microsimulation model.“, International Journal of Microsimulation, 6(1), 4-26. http://microsimulation.org/IJM/V6_1/2_IJM_6_1_Sutherland_Figari.pdf

Thank you for attention!

eesti statistika