LATVIJAS BANKA

EIROSISTĒMA

Circumstances defining the inequality of opportunity in Europe: what trees are telling us?

Ludmila Fadejeva, Krista Kalnbērziņa

16-17 March, 2023

8th EU-User Conference in Mannheim, Germany.

Motivation

► Inequality of opportunity directly linked to economic development potential.

► Assessing the ethically unacceptable part, or part defined by someone's circumstances, can explain the economic and institutional factors generating overall inequality.

► Identifying the circumstances that create the inequality of opportunity can help design targeted policies.
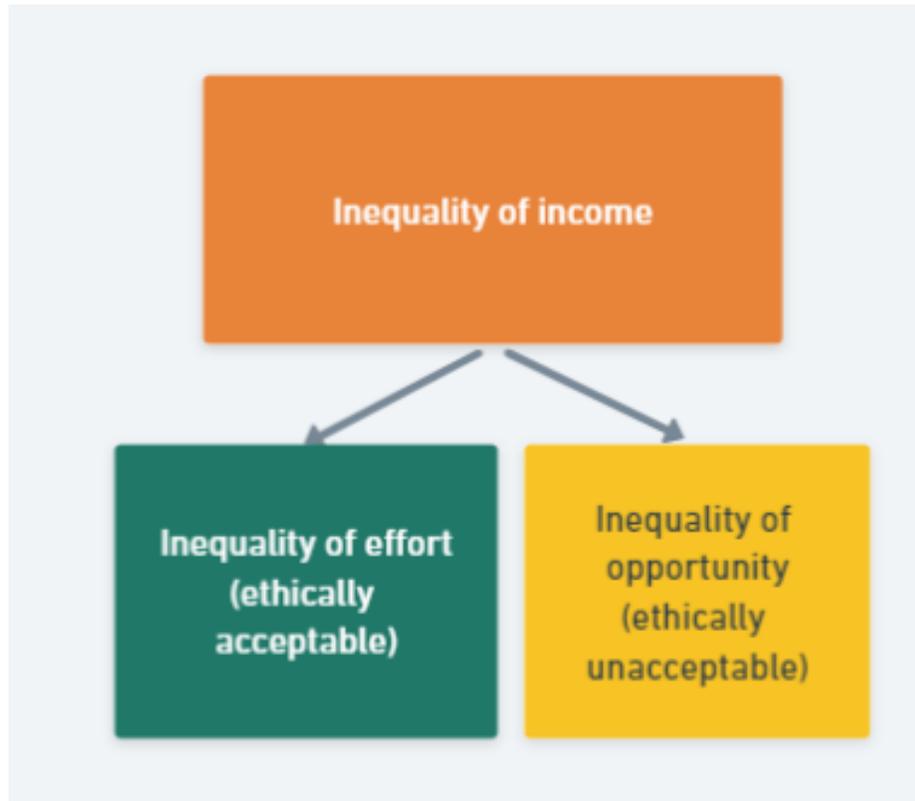
Literature – theoretical background

Philosophical debate and **inequality of opportunity (IO) in economic terms**.

► **John Roemer (1998)** defines the inequality of opportunity in economic terms. Divides inequality by effort, the factor the individual has control over, the circumstance that is given to the person.

► **Roemer and Trannoy (2016)** propose a full review of philosophical debate and economic models used to determine the inequality of opportunity.

Inequality of opportunity link to economic growth potential

► **Marrero and Rodriguez (2013)** find robust support for a negative relationship between IO and growth, and a positive relationship between inequality of effort and growth.

► **Ayar and Ebeke (2019)** suggest that omitting IO leads to misspecification, shedding light on why the empirical literature on income inequality and growth has been so inconclusive.

Literature – practical application

Measurement of IO stem from two premises (Fleurbaey and Peragine (2013)).

- ► **Ex-post** – there is equality of opportunity if all individuals exerting the same effort can achieve the same outcome.

- ► **Ex-ante** – evaluates individuals of the same circumstances and various levels of efforts. Equality of opportunity – all individuals face the same opportunity regardless of their circumstances.

► **Ramos and de gaer (2017)** concludes that the choice of ex- post or ex-ante approach can yield different country rankings.

► Vast literature propose various approaches to determine the level of IO

- ► **Checchi et al. 2010** distinguishes between ex- ante and ex- post and find that ex-ante equality correlate positively with public spending on education and ex - post to trade unions and fiscal redistribution.

- ► **Andreoli et al. 2021** analyse ex- ante inequality of opportunity by age cohorts. **Brunori et al. 2013** compares ex – ante inequality of opportunity in 41 countries.

- ► **Fleurbaey et al (2015)** calculate ex-ante IO by socioeconomic groups in Germany.

► The regression tree approach used in **Brunori et al. 2021** suggests that machine learning methods can improve estimations by lowering upward and downward biases.

Literature – data sources

EU-SILC

▶ Inequality of opportunity for European countries have been possible to measure using the EU Survey of Income and Living Conditions (EU-SILC) database.

▶ Among others:

   ▶ Marrero and Rodrıguez 2012

   ▶ Checchi et al. 2010

   ▶ Andreoli et al. 2021

   ▶ Brunori et al. 2021.

Other data sources:

▶ National surveys (among others)

   ▶ Bank of Italy Survey on Income and Wealth (Pace 2017)

   ▶ Spanish Survey on Income and Living Conditions (ECV) (Ayala et al 2021)

▶ Human Opportunity Index (HOI) together with Life in Transition Surveys (LiTS) (Abras et al. 2013)

Data

EU-Statistics on Income and Living Conditions (EU-SILC)

|	
Modules on Intergenerational transmission of disadvantages	2005, 2011, 2019
Questions are comparable (some questions are added or removed) (~21 question)	2011, 2019
Countries for which both 2011 and 2019 questions are available	AT, BE, BG, CH, CY, CZ, DK, EE, EL, ES, FI, FR, HR, HU, IE, IT, LT, LU, LV, MT, NL, NO, PL, PT, RO, SE, SI, SK
Sample	Age 26 – 40, employed, at least 3 months as full-time employee
Variable of interest	employed income per hour

Sample age
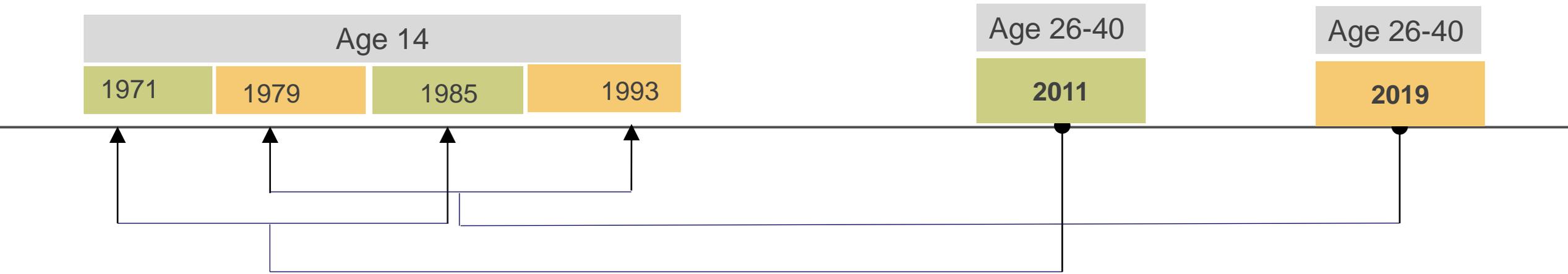
Circumstances (at 14):

- ► Number of adults in HH
- ► Number of children in HH
- ► Financial situation of the HH
- ► Farther/Mother
  - citizenship
  - highest level of education
  - activity status
  - managerial position
  - main occupation
  - year of birth
  - country of birth

- ► Ability to make ends meet **
- ► Tenancy status (owner/tenant)
- ► Degree of urbanization * (large city/ town/city)
- ► Basic school needs *
- ► Having meal with meat/fish daily *
- ► One-week annual holiday away from home*
- ► Type of household (private/collective) *

Other circumstances:

- ► Sex

Note: * - new question in 2019 EU-SILC, ** - question present in 2011 EU-SILC only

Examples of circumstances determining future income (EU country average)
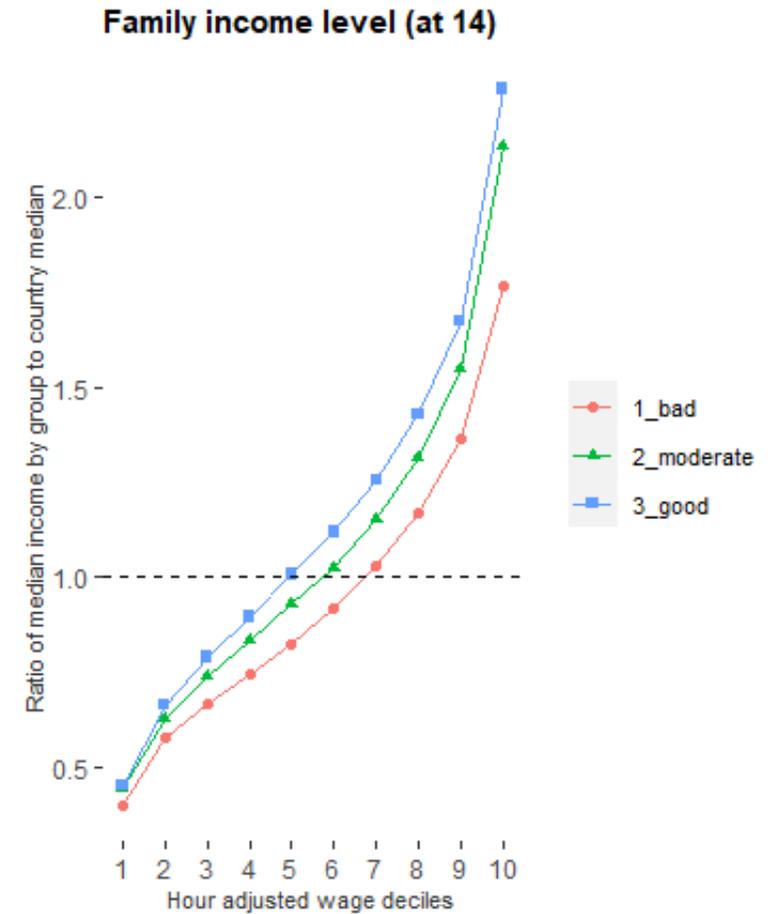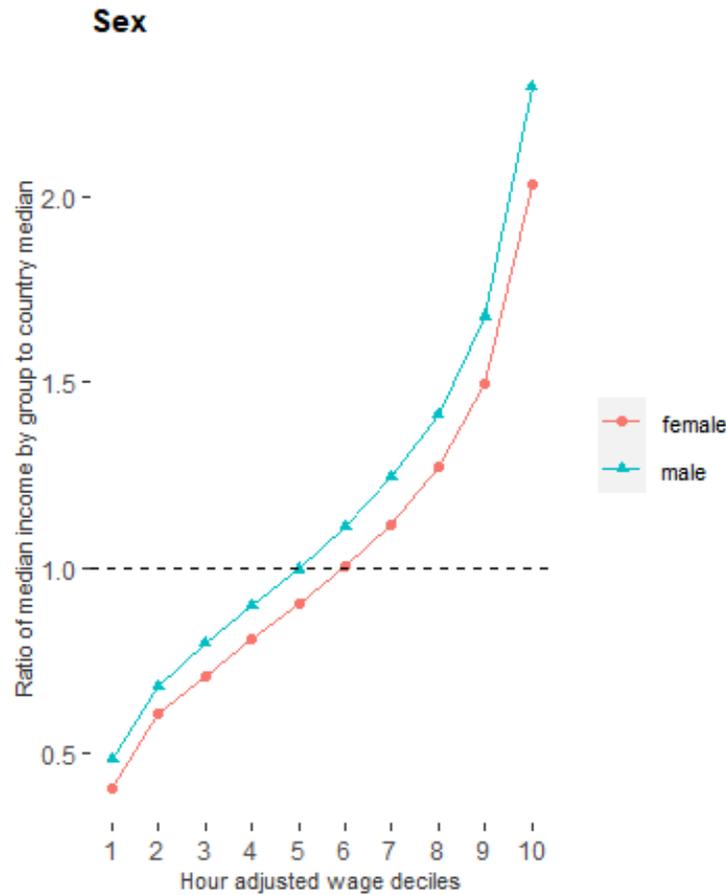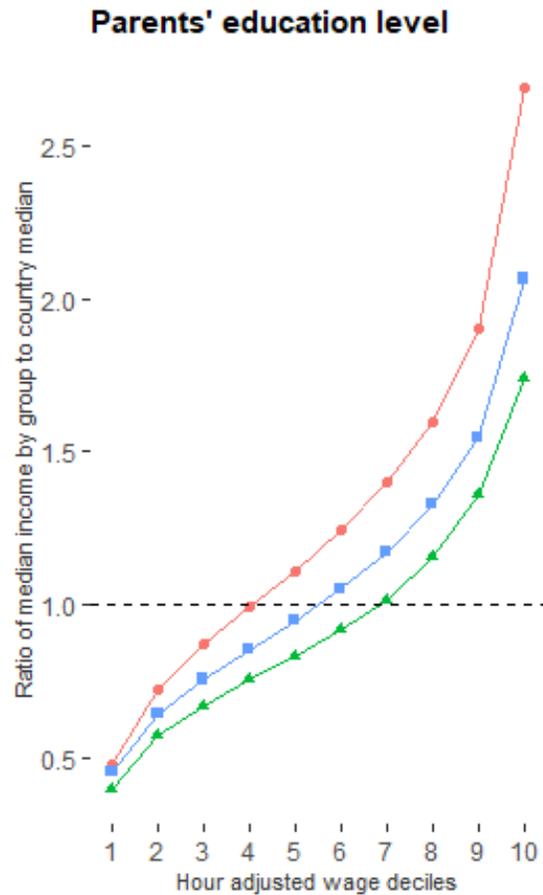
Data: EU-SILC, 2019

Note: for each country data divided into groups by circumstance, then ranged and deciles estimated (for each group). EU average is estimated as simple average across countries.

Methodology: how to estimate IO?

Problem points:

Y(total) = f(Circumstances, Effort)

► Effort is often unobservable

► Set of circumstances is limited to few categories

► Non-linear relations between circumstances and income (multiple interaction terms possible)

► Interrelations between circumstances and effort

Measurement approaches:

► Ex-Ante (circumstances known)

► Ex-Post (effort known)

► Non-parametric (means by grid of circumstances)

► Parametric (regressions with X(circumstances))

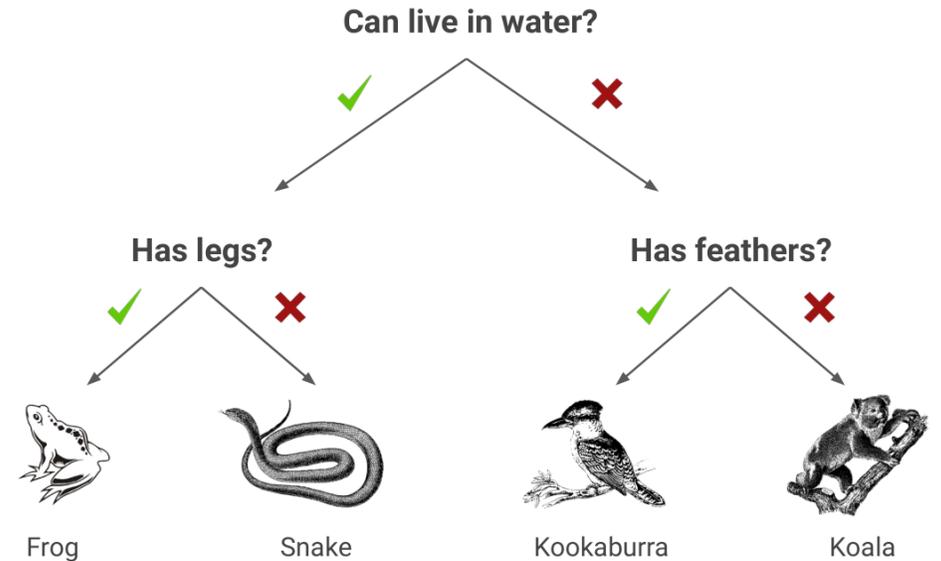► Trees (with circumstances defining terminal nodes/ leafs)

Why trees?   Non-linear relations are accounted for, immune to outliers, does not rely on underlying distribution or functional form
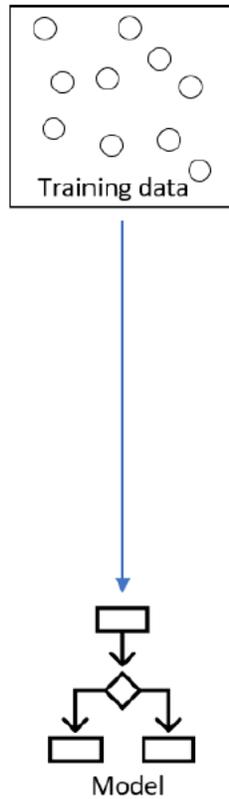
Trees (basic idea)

Steps:

► Divide sample into training (largest part) and testing

► Choose type: classification trees (for 1-0 problems) or regression trees (for level prediction problems)

► Choose measurement: for classification trees – confusion matrix; for regression trees - RMSE, MAE.

► Make sure that your predictor works equally well both in training and sample (no overfitting).

► Choose tuning parameters (differ by type of tree), e.g., depth of the tree

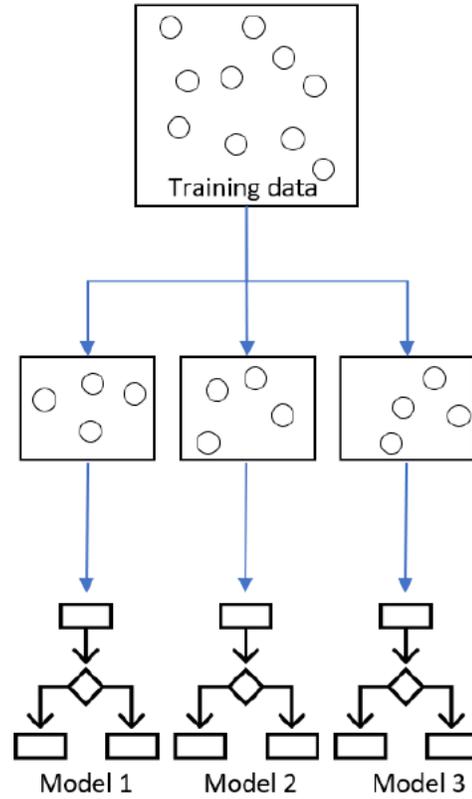► Extras: look for best tuning parameters over grid of values; use folds to repeat exercise for various subsamples
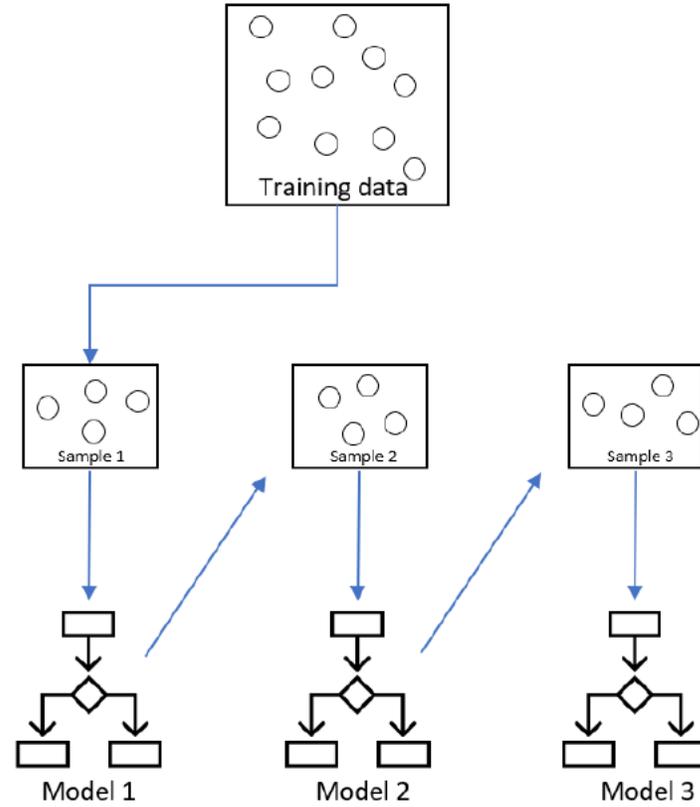

Can live in water?
✓ ✗
Has legs?   **Has feathers?**
✓ ✗   ✓ ✗
Frog   Snake   Kookaburra   Koala

Trees types

MACHINE LEARNING WITH TREE-BASED MODELS IN R

# Boosted trees (application to IO)

Choices made:

► Variables:

  o Most circumstances have ordered values, e.g., level of education – low, medium, high.

  o We have explored 2 options - (a) use trees with level circumstances, (b) created binary variables for each level of circumstance.

  o Many variables are about parents, but many HH had only single parent in childhood. To preserve all households, we create new household level variable, such as highest level of parents' education in family, etc.

► Tuning:

  o For each country/year we tuned over size of testing sample (0.55-0.75) and tree depth (3-10).

  o Number of trees is set at 10, other parameters are set at default values.

  o We use RMSE to pick the best specifications.

Training and tuning results

RMSE (training vs testing)
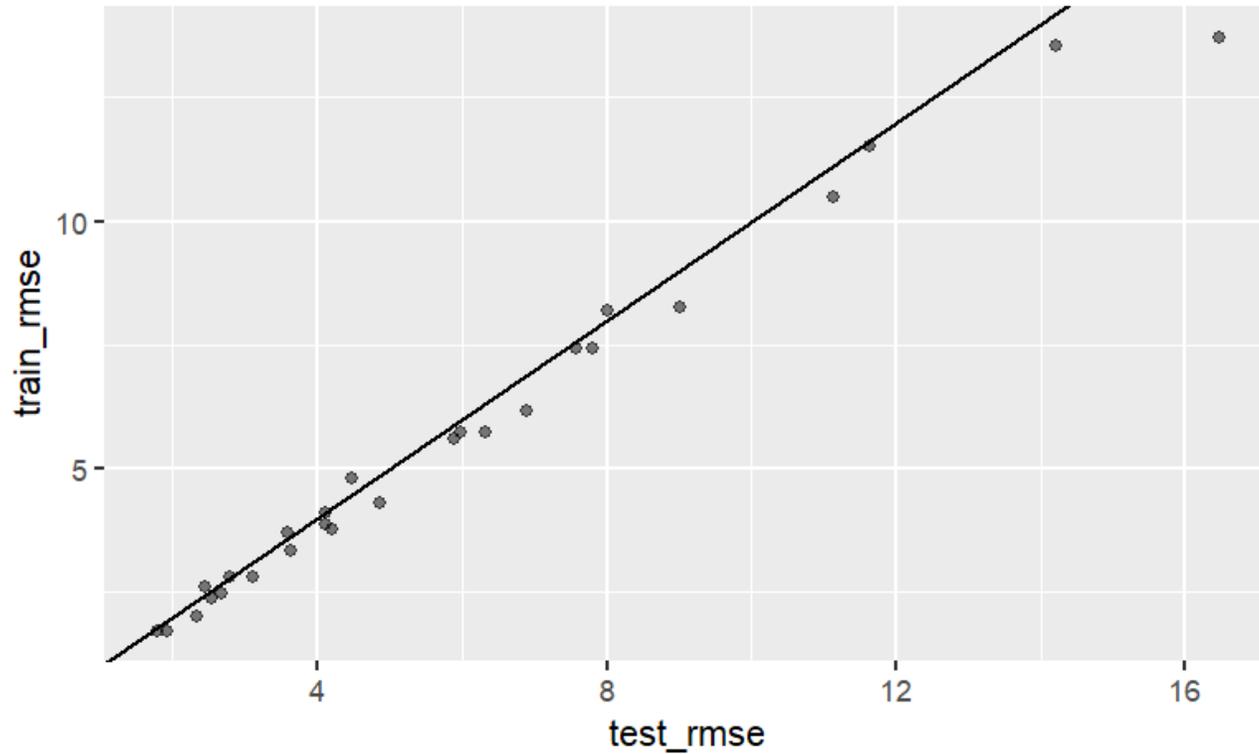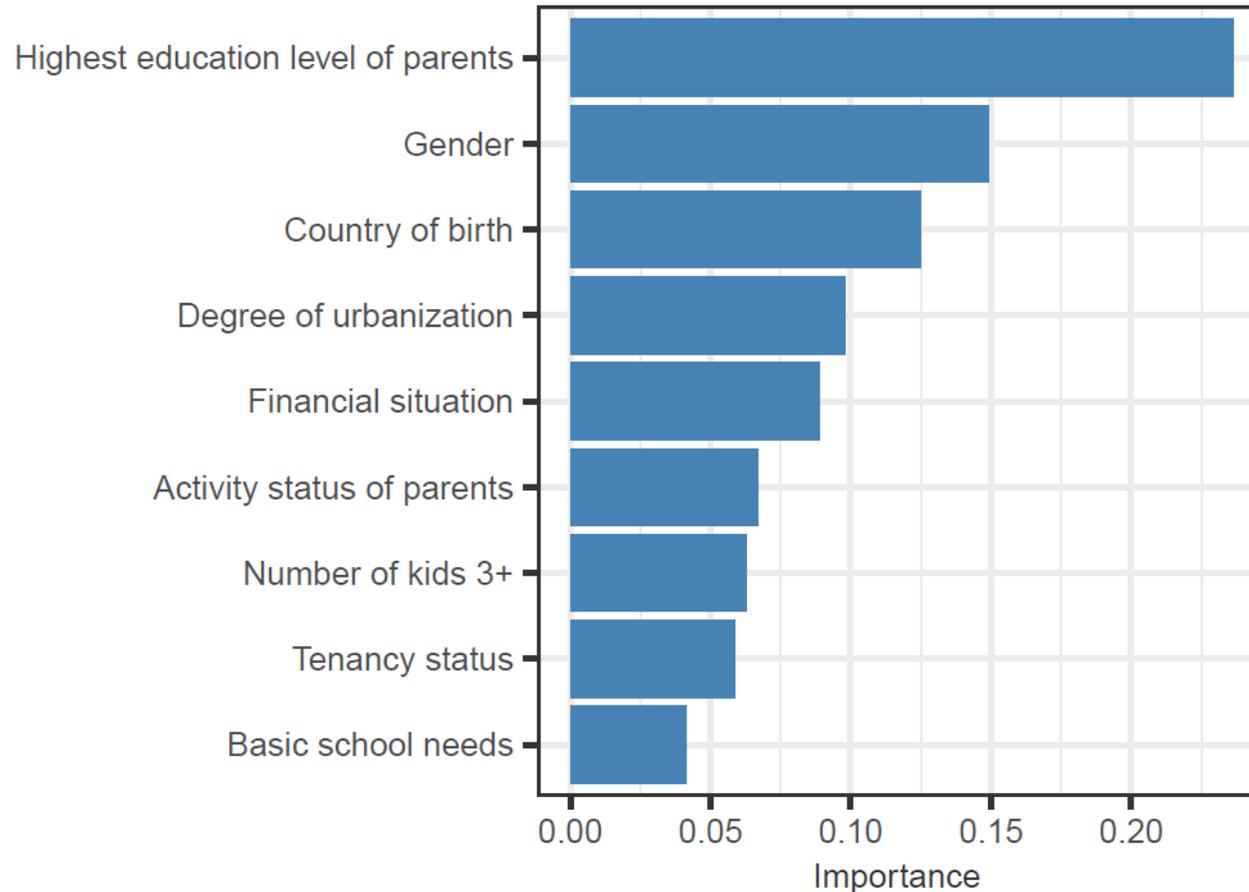

Preferred tree debts:

- Majority 3 (some 4-5)

Sample size :

- Majority 0.65-0.75

EU main factors determining difference in hourly wage (generating inequality of opportunity)

Importance factor (average EU in 2019)

Policy areas concerned:

► Access to education

► Accessible childcare

► Gender equality

► Decent employment

► Integration policy

Source: authors' estimation. Boosted tree model specification with binary variables.

Top 5 inequality viewed by circumstance variables

Lithuania, Latvia

- Education level of a parent
- Gender
- Financial situation of a family

Bulgaria

- Education level of a parent
- Degree of urbanization
- Gender

Spain

- Parent country of birth
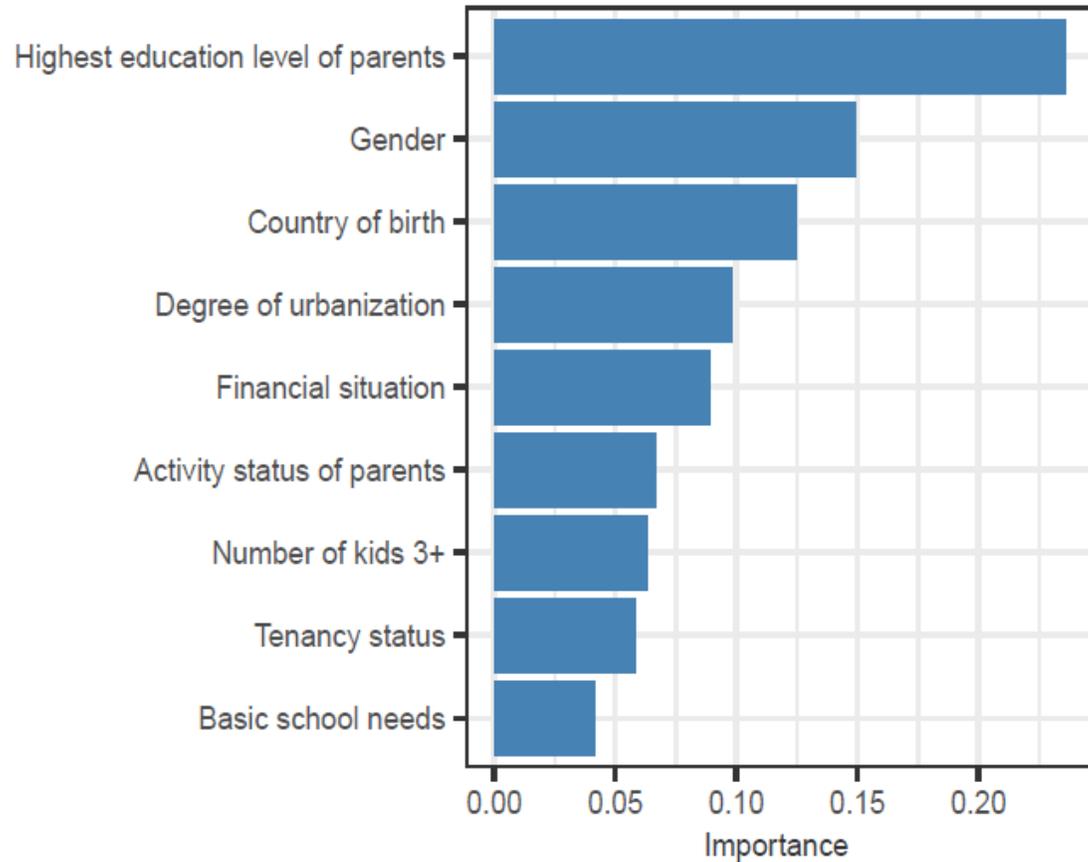- Education level of a parent
- Gender

Greece

- Education level of a parent
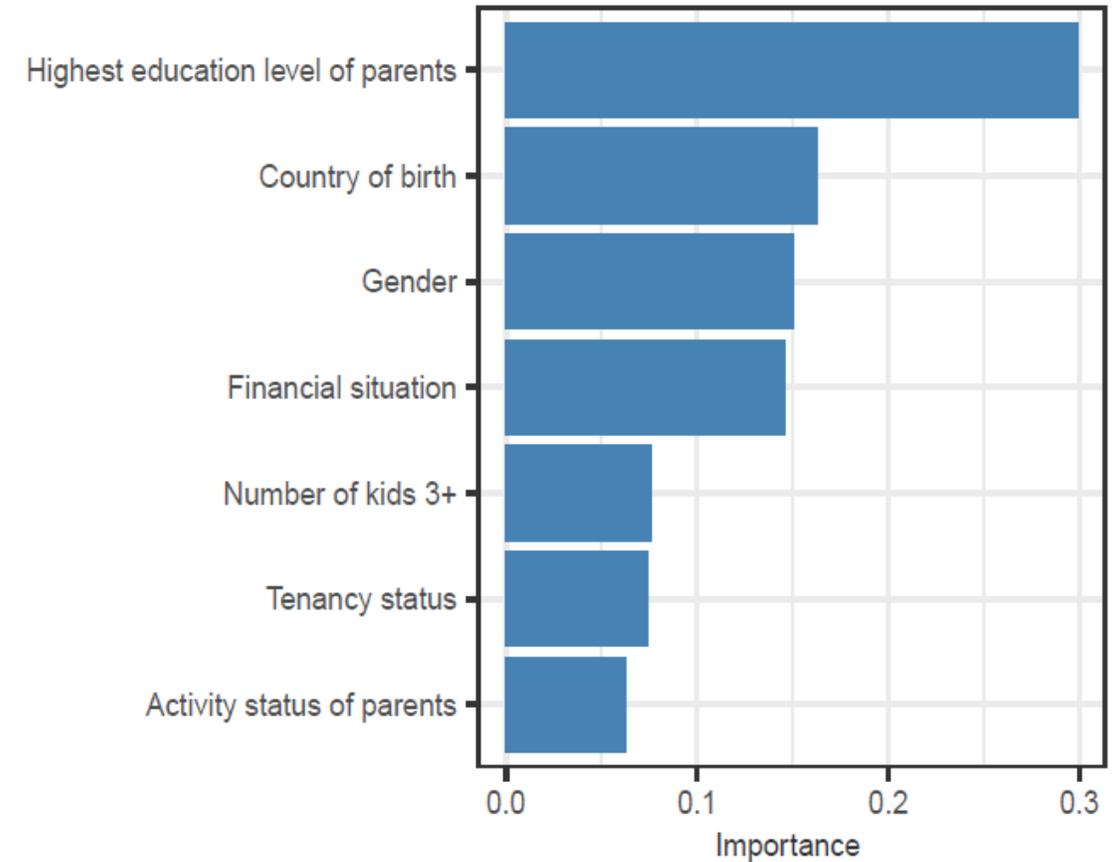- Gender
- Financial situation

Estonia

- Education level of a parent
- Gender
- Degree of urbanisation

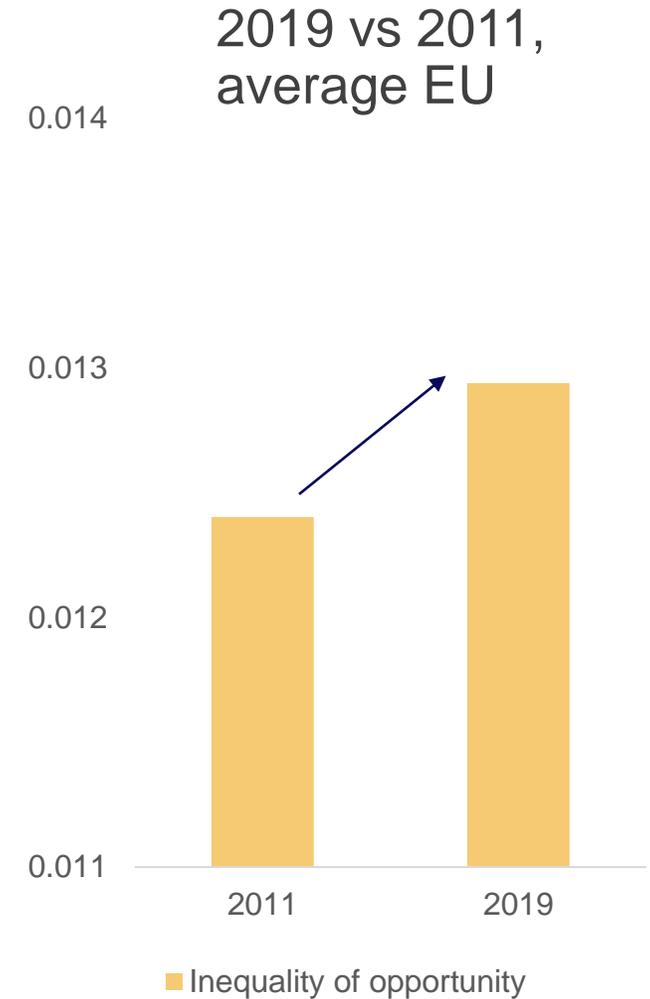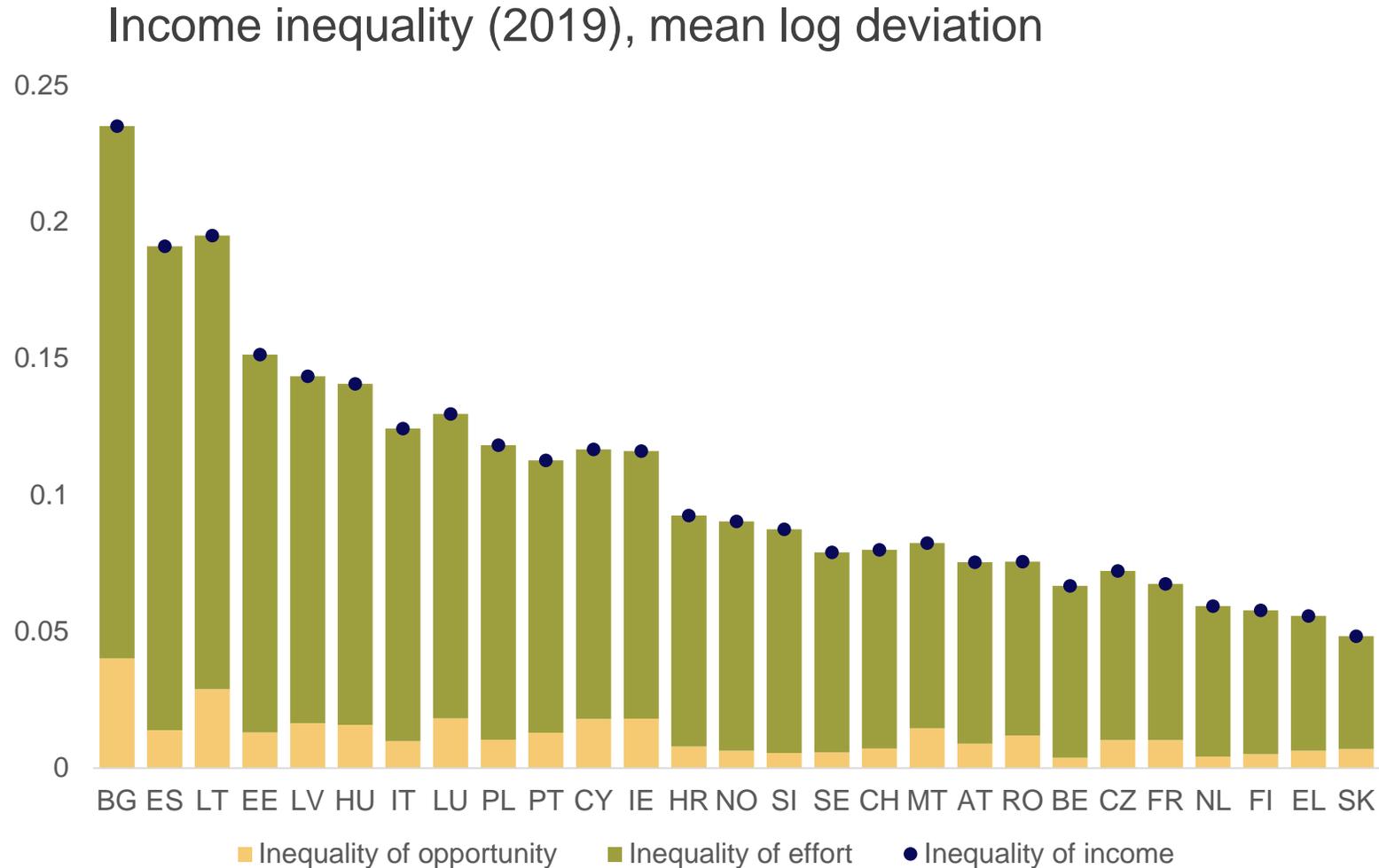# Average EU level importance factors: 2011 vs 2019

(a) 2019

(b) 2011

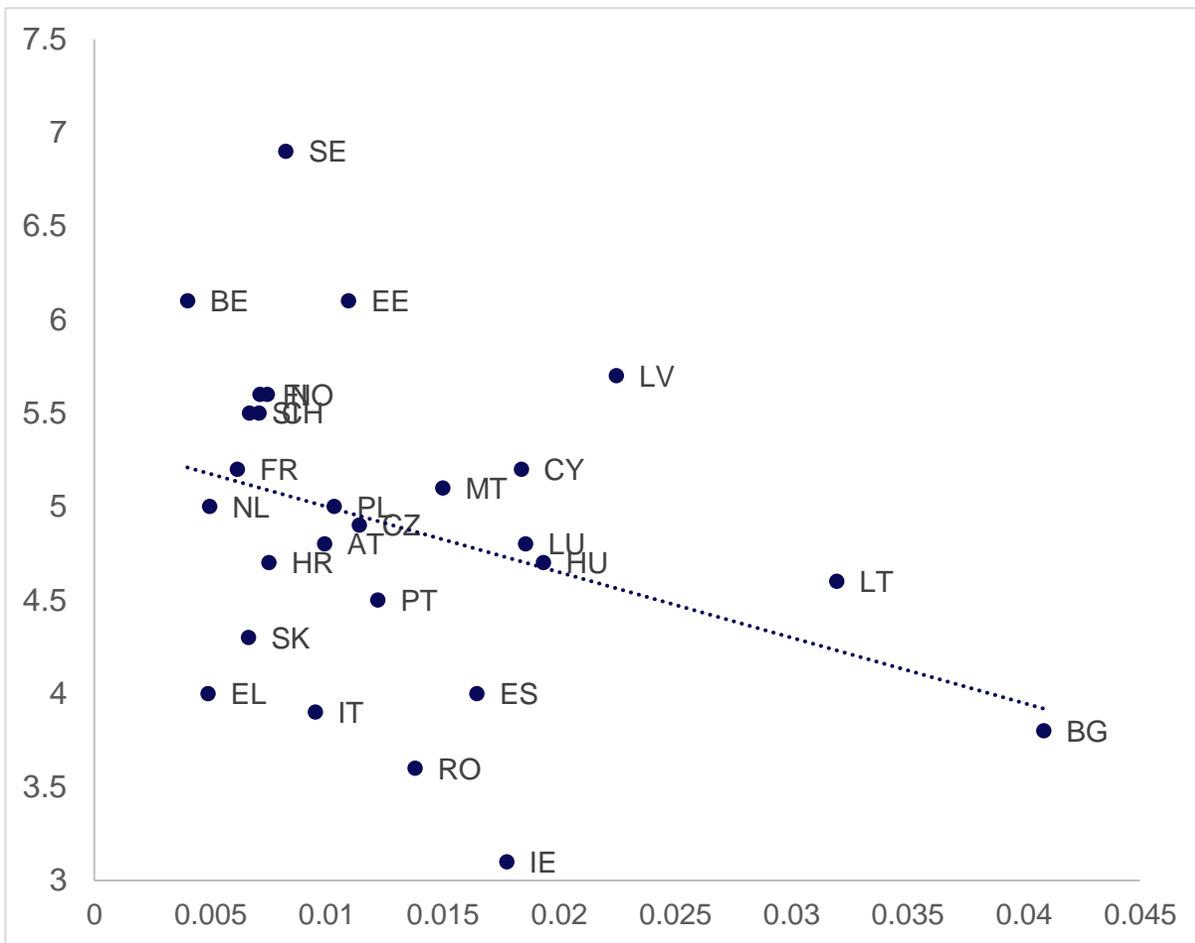# Share of inequality of income due to IO range from 6% - 18%

Income inequality (2019), mean log deviation

2019 vs 2011, average EU

Legend: Inequality of opportunity · Inequality of effort · Inequality of income

Inequality of opportunity

20

IO negative correlation to government social expenditure in 2019

Social protection/GPD and IO
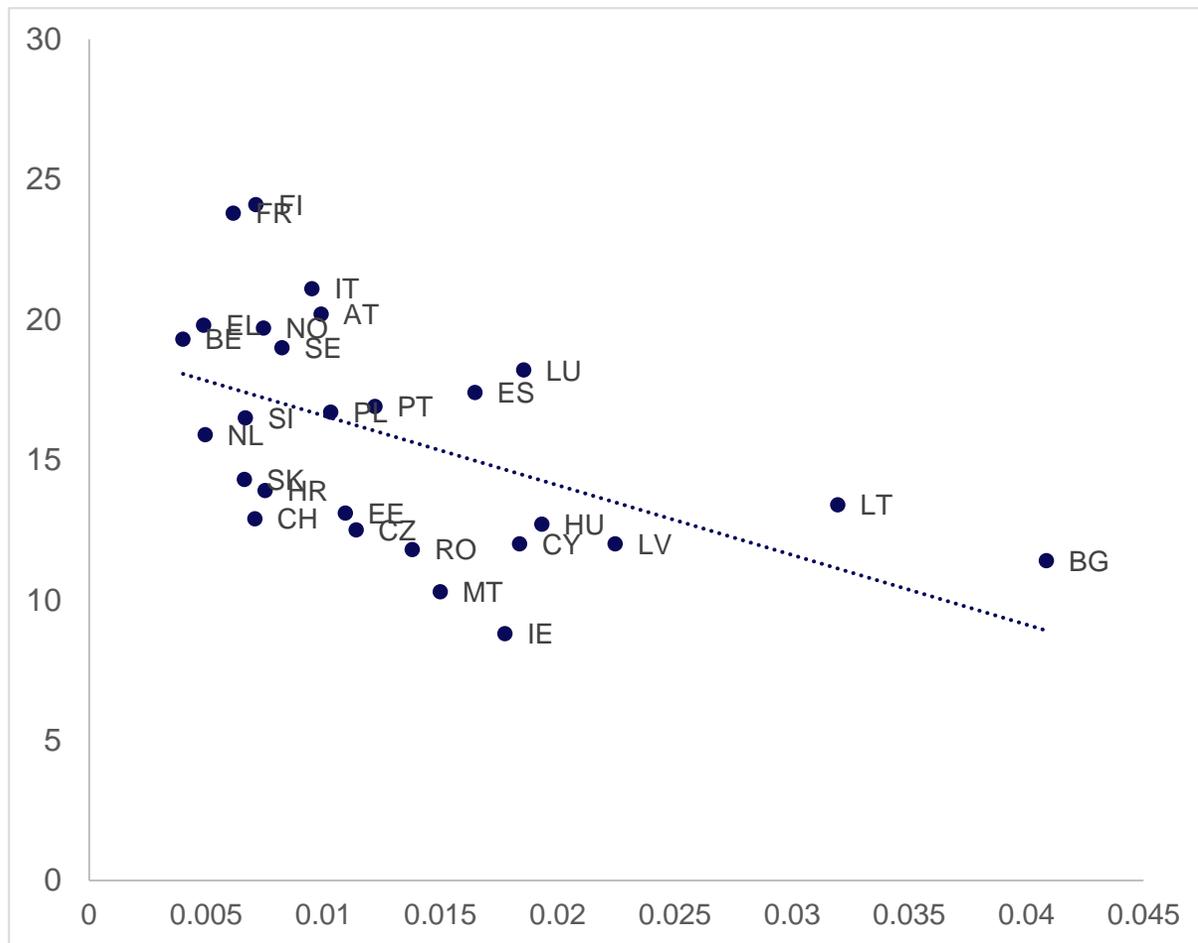
Education expenditure/GDP and IO

Source: Authors` own calculations.

21

Conclusions

► Inequality of opportunity has slightly increased in 2019 compared to 2011.

► Share of inequality of opportunity in inequality of income and range between 6%-18%.

► Main childhood circumstances that influence the income today are education of parents, children/adult ratio in the family and gender.

► In order to equalise opportunities for next generations it is important to adjust policies aiming increase education accessibility, appeal to those that come from less educated households, and availability of lifelong learning. Gender equality policies. Childcare and protection policies.

Things to do

1. Robustness check (variable specification and selection)

2. Extend tuning exercise (more tuned parameters)

3. Use wider set of estimation methods for comparison

4. Policy implications