

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

Ludmila Fadejeva¹, Krista Kalnberzina²

Key words: inequality of opportunity, intergenerational transmission of disadvantages, boosted trees; EU-SILC

Inequality of opportunity is directly linked to economic development potential (i.e., Marrero and Rodriguez, 2013). Assessing the ethically unacceptable part, or part defined by someone's circumstances, can explain the economic and institutional factors generating overall inequality, as proposed by Roemer (1998). The module on intergenerational transmission of disadvantages in EU-SILC allows exploring this topic (see for example Marrero and Rodriguez (2012), Checchi et.al. (2010), Brunori et.al. (2013))³. Most previous studies focus on estimation of the inequality of opportunity measures distinguishing between ex-ante/ ex-post measures and non-parametric and parametric methods (e.g., Ramos and Van de gaer (2017)). Much smaller focus is on the factors determining the inequality of opportunity (Han (2022)).

The research questions of our study are: Which are the main circumstances generating inequality of income in adulthood? Do they differ by country? Which is the share of inequality of income defined by one's circumstances? Do these results change in between 2011 and 2019? What is association between government social expenditure and inequality of opportunities?

Traditionally the group of methods allowing evaluating the importance of circumstances in inequality of opportunity are parametric methods (e.g., Ferreira, F. H. and Gignoux, J. (2011)). Traditional regression or instrumental regression methods suffer from the estimation biases, which are hard to fix due to limited number of explanatory variable available in the survey. Brunori et al. (2021) shows that machine learning methods, in particular regression tree approach, improves estimations by lowering those upward and downward biases. We follow this approach and apply two tree methods - random forest and boosting - proven to provide good explanatory power in the class of tree models to the dataset.

So far, there were three waves of EU-SILC module on intergenerational transmission of disadvantages (2005, 2011, 2019). Due to better question comparability and larger sample of countries participating in the last two waves (28 European countries⁴) we focus on 2011 and 2019. The novelty of our paper stems from three factors – we use the latest data available, focus on circumstances determining the inequality of opportunity rather than level, and use novel estimation methods.

We show that the overall inequality of opportunity in Europe has slightly increased in 2019 as compared to 2011. The share of inequality of opportunity in inequality of income is heterogenous between different European countries and range between 6%-18%. Three main childhood factors that influence the income today are education of parents, children/adult ratio in the family and gender. The ranking of those factors differs between countries; however, these top factors remain of the highest importance in both 2011 and 2019 at the aggregate level. We also show that there is positive association between expenditures on social protection to GDP or education expenditures to GDP and inequality of opportunity.

¹ Latvijas Banka, Monetary Policy Department

² Latvijas Banka, Monetary Policy Department

³ Other possible data sources to investigate inequality of opportunity are national surveys, e.g., Bank of Italy Survey on Income and Wealth (Pace 2017), Spanish Survey on Income and Living Conditions (ECV) (Ayala et al 2021), and Human Opportunity Index (HOI) together with Life in Transition Surveys (LiTS) (Abrás et al. 2013).

⁴ 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