DOI: 10.22363/09669-2019-666-673

MORE EQUAL AND LESS EQUAL AT THE SAME TIME? MEASURING INEQUALITY IN EDUCATIONAL ACHIEVEMENT OF 15-YEAR OLDS IN 37 COUNTRIES

Gwyther Rees[1], ***Anna Gromada***[2], ***Yekaterina Chzhen***[3]

[1]The UNICEF Office of Research – Innocenti, Florence, Italy,
e-mail: grees@unicef.org.
[2]The UNICEF Office of Research – Innocenti, Florence, Italy and Polish Academy of Sciences, Warsaw, Poland, e-mail: agromada@unicef.org
ORCID ID: 0000-0001-8135-1424
[3]The UNICEF Office of Research – Innocenti, Florence, Italy,
e-mail; ychzhen@unicef.org

Abstract. The article aims to explore how different approaches to the conceptualisation and measurement of educational inequalities affect the conclusions that are drawn from comparisons between countries and over time. It contains new analysis based on the most recent waves of the Programme for International Student Assessment (PISA). It shows that different measures of inequality lead to diverging, and sometimes even contradictory, conclusions about which countries are faring better and worse in terms of educational equality compared to other countries, and over time. We focus on countries that are a member of the European Union (EU) and/or OECD. The study produces nationally representative and cross-nationally comparable data on schoolchildren's skills and knowledge in reading.

According the Pearson correlations between each pair of inequality measures across 37 countries, using data from PISA 2015 all measures of inequality of outcome are strongly related to one another - with correlations ranging from 0.92 to 1.00. Measures of inequality of opportunity show somewhat weaker relationships, with correlations ranging from 0.56 to 0.83. The limited range (P90-P10) is a good 'representative' measure of inequality of outcome for international comparisons because it is easy to communicate to non-specialist audiences (see UNICEF Office of Research 2018). It shows perfect correlation with standard deviation and very high correlation (0.93***) with the Gini Index. P90-P10 will be subsequently compared with three measures of inequality of opportunity.

The complex relationship between the inequality of outcome and inequality of opportunity will be discussed by comparing Iceland and Hungary. The two countries find themselves on the two extremes of the inequality of opportunity spectrum. Iceland is the single most "equal" country using ESCS R-squared and the third most equal using the gradient. As such, it

emerges as a country with the single weakest influence of economic, social and cultural status on children's reading skills. Yet, Iceland's ranking goes down radically to the 24th rank when we opt for the R-squared which takes into account factors outside of children's control that are not directly linked to their parents (i.e. school location, migration, language and gender).
By contrast, Hungary emerges as one of the three most "unequal" countries using all three measurements. It is the single most unequal using F-G and ESCS R-squared and the third most unequal using the ESCS gradient. The record 35% of reading score variance can be attributed to factors outside children's control, including 22 percentage points attributable solely to ESCS status – making for an especially strong link between the family status and educational outcomes. However, the most "equal" Iceland and the most "unequal" Hungary turn out to have almost identical inequality of outcome regardless of the measurement method. They show the standard deviation of 99 and 97 respectively, Gini of 0.10-0.11 and P90-P10 of 255-256 points. Furthermore, these outcomes place them in the bottom half of the equality of outcome ranking among the analysed countries.
Keywords: inequality of opportunity, inequality of outcome, educational achievement, reading, PISA

Introduction

There is a growing recognition among international organizations, scholars and policymakers that education systems should focus not only on raising standards but also on reducing inequalities, but there is far less consensus on what this means in practice. Previous studies have tended to take either the inequality of outcome or inequality of opportunity approach. Yet, comparisons of their relative merits and implications in the field of education are rare.

Purpose of the study

This article aims to explore how different approaches to the conceptualisation and measurement of educational inequalities affect the conclusions that are drawn from comparisons between countries and over time. It contains new analysis based on the most recent waves of the Programme for International Student Assessment (PISA). It shows that different measures of inequality lead to diverging, and sometimes even contradictory, conclusions about which countries are faring better and worse in terms of educational equality compared to other countries, and over time.
The paper makes a unique contribution to the study of comparative educational inequalities in three ways. First, it draws attention to how the conceptualisation and measurement of educational inequality can affect what we understand as 'inequality'. Second, it is the most comprehensive

and up-to-date analysis of different approaches to making international comparisons of inequality in educational achievement between countries and over time. Third, it shows how the choice of the measurement method can influence the messages that are extracted in different countries and therefore also can affect the policy implications that are drawn from these messages.

Methods

Data and variables

We focus on countries that are a member of the European Union (EU) and/or OECD. We analyse data from two recent rounds of the PISA study in 2009 and 2015. The study produces nationally representative and cross-nationally comparable data on schoolchildren's skills and knowledge in reading. PISA features a rich range of student and family socio-economic characteristics, includes three subject areas at once and covers a greater number of EU/OECD countries. PISA reading scores were standardized for the survey wave in the year 2000 so that at that point the test scores had an international mean of 500 points and a standard deviation of 100 points. Such standardization facilitates comparability across survey waves.

Measures of inequality

The article compares two approaches to the measurement of inequality – inequality of opportunity and inequality of outcome – while also testing the robustness of the cross-country comparative results for each approach using three measures for each of the concepts.

Inequality of outcome measures:

(1) The Gini Coefficient
(2) The restricted range of the reading score distribution (P90–P10)
(3) Standard Deviation

Inequality of opportunity measures:

(4) The slope of the ESCS gradient

The ESCS slope is the coefficient derived from a simple linear regression without controls of student reading scores onto a variable measuring the economic, social and cultural status.

(5) ESCS R-squared: the share of variance in reading achievement that can be explained by economic, social and cultural status.

(6) Ferreira and Gignoux R-squared (F-G): the share of variance in reading achievement that can be explained by gender, parental education, parental occupation, language spoken at home, migration status, number of books at home, number of cultural possessions at home, household wealth index and for location of school.

Results

Comparing inequality of outcomes and opportunities

Table 1 shows the Pearson correlations between each pair of inequality measures across 37 countries, using data from PISA 2015. All measures of inequality of outcome are strongly related to one another - with correlations ranging from 0.92 to 1.00. Measures of inequality of opportunity show somewhat weaker relationships, with correlations ranging from 0.56 to 0.83. The limited range (P90-P10) is a good 'representative' measure of inequality of outcome for international comparisons because it is easy to communicate to non-specialist audiences (see UNICEF Office of Research 2018). It shows perfect correlation with standard deviation and very high correlation (0.93***) with the Gini Index. P90-P10 will be subsequently compared with three measures of inequality of opportunity.

Table 1 - Correlation matrix of inequality scores

	Inequality of outcome			**Inequality of opportunity**		
	P90-P10	SD	Gini	ESCS gradient	ESCSR^2	F-G
P90-P10	.					
SD	1.00***	.				
Gini	0.93***	0.92***	.			
ESCS gradient	0.65***	0.63***	0.57***	.		
ESCS R^2	0.36*	0.32(ns)	0.45**	0.67***	.	
F-G	0.41*	0.37*	0.54***	0.56***	0.83***	.

Note: Correlations based on scores (not ranks) of 37 countries.
P-values: ns ($p > 0.05$) *($p \leq 0.05$), **($p \leq 0.01$), ***($p \leq 0.001$).
Source: PISA, 2015.

Second, we compare the inequality of outcome with inequality of opportunity conceptualized as the ESCS gradient. Linear regressions were run with no controls for each country, with the reading score as the dependent variable and the socio-economic status scale as the independent variable. The ESCS status has a statistically significant association with reading scores (with p-values less than 0.001) in all countries. The gradient ranges from 26 in Spain to 59 in Luxembourg. The gradient has a correlation R=0.65 with the P90-P10 measure (Figure 1).

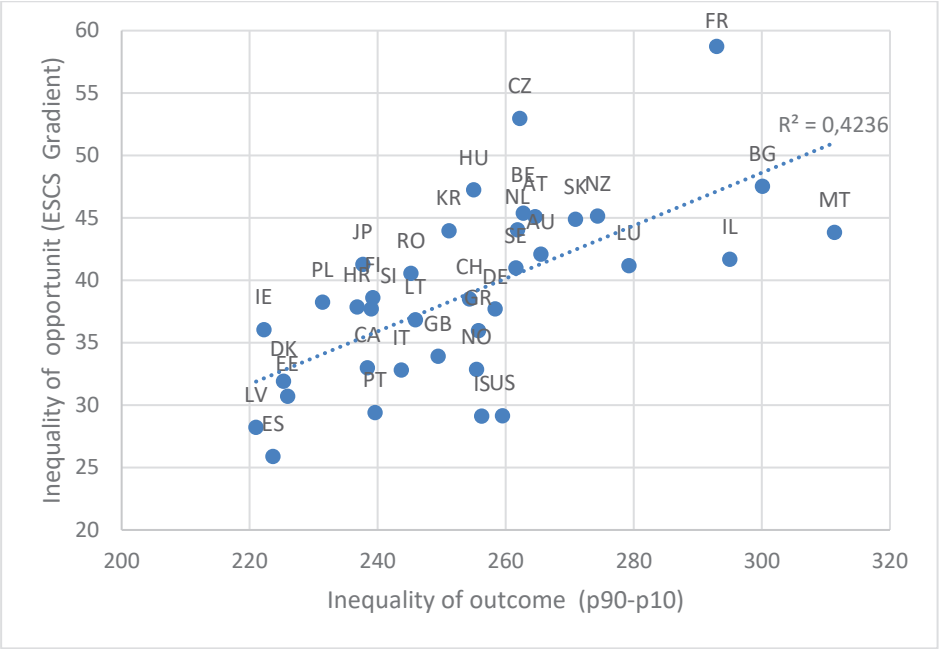

Figure 1 - Inequality of outcome versus inequality of opportunity measured as the ESCS gradient

Source: PISA, 2015.

Third, we compare the inequality of outcome measure with inequality of opportunity conceptualized as ESCS R-squared, or the share of variance in reading achievement that can be explained by economic, social and cultural status. It explains on average 12% of the gaps in reading skills: ranging from 5% in Iceland to 22% in Hungary. It has a correlation of 0.36 with inequality of outcome (Figure 2).

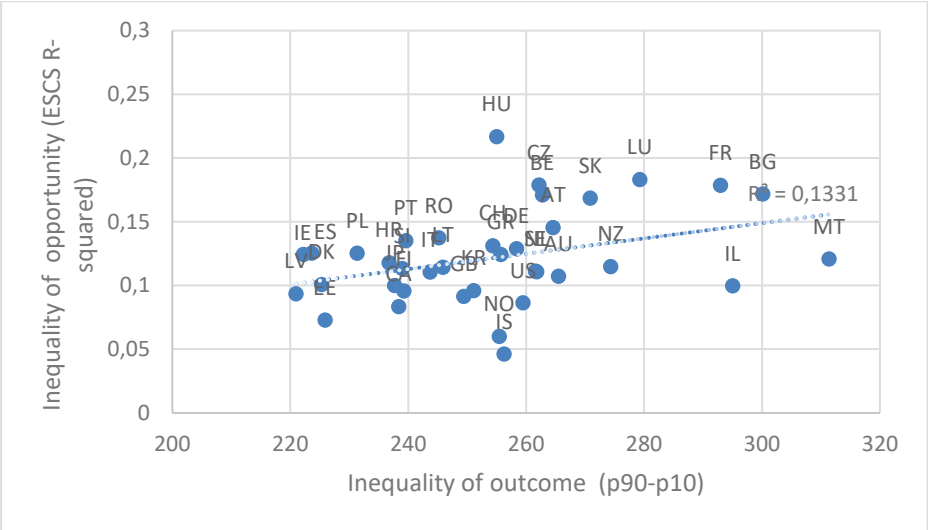

Figure 2 - Inequality of outcome versus inequality of opportunity measured as the ESCS R-squared
Source: PISA, 2015.

Fourth, we compare the inequality of outcome with inequality of opportunity measured by the overall R-squared statistic from a linear regression (F-G). On average, the independent variables explain around 22 per cent of the variation in test scores. This ranges from 14 per cent in Japan to 35 per cent in Hungary. When juxtaposed with the inequality of outcome, it shows the correlation of 0.41 (Figure 3).

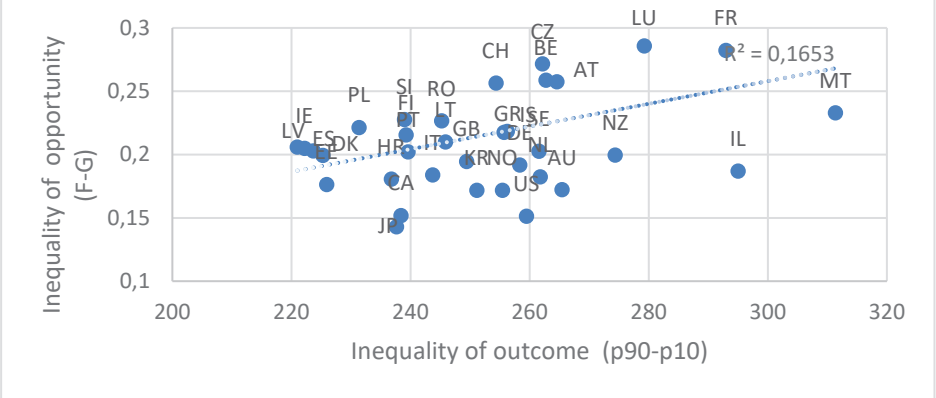

Figure 3 - Inequality of outcome versus inequality of opportunity measured as overall R-squared (F-G)
Source: PISA, 2015.

The complex relationship between the inequality of outcome and inequality of opportunity will be discussed by comparing Iceland and Hungary. The two countries find themselves on the two extremes of the inequality of

opportunity spectrum. Iceland is the single most "equal" country using ESCS R-squared and the third most equal using the gradient. As such, it emerges as a country with the single weakest influence of economic, social and cultural status on children's reading skills. Yet, Iceland's ranking goes down radically to the 24th rank when we opt for the R-squared which takes into account factors outside of children's control that are not directly linked to their parents (i.e. school location, migration, language and gender).

By contrast, Hungary emerges as one of the three most "unequal" countries using all three measurements. It is the single most unequal using F-G and ESCS R-squared and the third most unequal using the ESCS gradient. The record 35% of reading score variance can be attributed to factors outside children's control, including 22 percentage points attributable solely to ESCS status – making for an especially strong link between the family status and educational outcomes. However, the most "equal" Iceland and the most "unequal" Hungary turn out to have almost identical inequality of outcome regardless of the measurement method. They show the standard deviation of 99 and 97 respectively, Gini of 0.10-0.11 and P90-P10 of 255-256 points. Furthermore, these outcomes place them in the bottom half of the equality of outcome ranking among the analysed countries.

Discussion

Our results suggest that indicators of variation in educational outcomes are more robust to the types of problems that affect international comparisons of educational achievement than common measures of inequality of opportunity. Across the 37 countries analyzed, the gap between the 90th and 10th percentiles (P90-P10) is highly correlated with two other popular indicators of inequality of outcome - the Gini index and the standard deviation of the scores. It does not matter much where the cut-off is made along the distribution: the P90-P10 gap is highly correlated with other range-based indicators (i.e. P95-P5 and P75-P25). Most importantly, if the educational outcome scale itself is fit for international comparisons, so would be the indicator that summarizes the degree of inequality in its distribution.

In contrast, as measures of inequality of opportunity exploit the relationship between the outcome and the students' circumstances, the potential for bias is greater. Measurement error in the predictor variables (i.e. circumstances) would bias the estimate of their association with the outcome. Moreover, all the circumstances need to mean the same thing in all the countries for valid international comparisons of inequality of opportunity. This is not necessarily true across the countries participating in PISA. For example, Põder, Lauri, and Veski (2017, 677) raise questions about the cross-cultural comparability of items such as books at home: "In large-number inter-

country studies, it is difficult to argue that in different cultures or living conditions 'books' have the same effect on PISA scores; or, put simply, individuals in different cultures do not share the same aspiration to books." Such issues will be magnified as the diversity of countries included in the analysis increases.

We show that the lessons that could be drawn from country rankings of inequality of opportunity depend strongly on which circumstances are included in the measure. Whether we only use the socio-economic and cultural background of the family or also add the student's gender and the location of the school makes a visible difference to the results. Yet even the full suite of predictors does not explain more than one-third of the variation in the reading scores. If we follow Ferreira and Gignoux (2014) to interpret this as the share of variation due to the circumstances outside of children's control, it will appear that at least two-thirds of the variation in reading test scores lies *within* children's control (as a measure of their *effort*). Yet a lot of the unexplained variation is likely due to unmeasured circumstances, pure luck and measurement error.

Additionally, countries that appear on the opposite ends of the "opportunity" spectrum – such as the most "equal" Iceland and the most "unequal" Hungary turn out to have almost identical inequality of outcome regardless of the measurement method. This raises the normative question: is it more important to have small achievement gaps among children within a country or to have gaps of any size but easily attributable to circumstances we can measure? In terms of actual reading outcomes, Iceland is not more equal than Hungary. However, it benefits from the fact that we do not understand well the sources of its inequalities so that they fall into unexplained part of variance ($1\text{-}R^2$). By contrast, Hungary's inequalities can be easily framed in classic sociological theories with their focus on parental status – a research tradition that over the years has translated into sophisticated variables that are included into the array of "circumstances" and, consequently, into the explained part of variance (R^2).

References

Ferreira, F.H.G. and Gignoux, J. 2014. 'The Measurement of Educational Inequality: Achievement and Opportunity', World Bank Economic Review, 28 (2): 210–246.

Põder, K., Lauri, T. and Veski, A. 2017. Does School Admission by Zoning Affect Educational Inequality? A Study of Family Background Effect in Estonia, Finland, and Sweden, *Scandinavian Journal of Educational Research*, 61(6): 668–688.