unicef
Office of Research-Innocenti

Education for All? Measuring inequality of educational outcomes among 15-year-olds across 39 industrialized nations

Zlata Bruckauf and Yekaterina Chzhen

Office of Research - Innocenti Working Paper
WP-2016-08 | April 2016

INNOCENTI WORKING PAPERS

UNICEF Office of Research Working Papers are intended to disseminate initial research contributions within the programme of work, addressing social, economic and institutional aspects of the realization of the human rights of children.

The findings, interpretations and conclusions expressed in this paper are those of the authors and do not necessarily reflect the policies or views of UNICEF.

This paper has been peer reviewed both externally and within UNICEF.

The text has not been edited to official publications standards and UNICEF accepts no responsibility for errors.

For readers wishing to cite this document we suggest the following form:

Bruckauf, Z. and Y. Chzhen (2016). Education for All? Measuring inequality of educational outcomes among 15-year-olds across 39 industrialized nations, *Innocenti Working Paper* No.2016-08, UNICEF Office of Research, Florence.

ISSN: 1014-7837

THE UNICEF OFFICE OF RESEARCH – INNOCENTI

In 1988 the United Nations Children's Fund (UNICEF) established a research centre to support its advocacy for children worldwide and to identify and research current and future areas of UNICEF's work. The prime objectives of the Office of Research are to improve international understanding of issues relating to children's rights and to help facilitate full implementation of the Convention on the Rights of the Child in developing, middle-income and industrialized countries.

The Office aims to set out a comprehensive framework for research and knowledge within the organization, in support of its global programmes and policies. Through strengthening research partnerships with leading academic institutions and development networks in both the North and South, the Office seeks to leverage additional resources and influence in support of efforts towards policy reform in favour of children.

Publications produced by the Office are contributions to a global debate on children and child rights issues and include a wide range of opinions. For that reason, some publications may not necessarily reflect UNICEF policies or approaches on some topics. The views expressed are those of the authors and/or editors and are published in order to stimulate further dialogue on child rights.

The Office collaborates with its host institution in Florence, the Istituto degli Innocenti, in selected areas of work. Core funding is provided by the Government of Italy, while financial support for specific projects is also provided by other governments, international institutions and private sources, including UNICEF National Committees.

For further information and to download or order this and other publications, please visit the website at www.unicef-irc.org.

Correspondence should be addressed to:

UNICEF Office of Research - Innocenti
Piazza SS. Annunziata, 12
50122 Florence, Italy
Tel: (+39) 055 20 330
Fax: (+39) 055 2033 220
florence@unicef.org
www.unicef-irc.org
@UNICEFInnocenti
facebook.com/UnicefOfficeofResearchInnocenti

EDUCATION FOR ALL? MEASURING INEQUALITY OF EDUCATIONAL OUTCOMES AMONG 15-YEAR-OLDS ACROSS 39 INDUSTRIALIZED NATIONS

Zlata Bruckauf[1] (zbruckauf@unicef.org) and Yekaterina Chzhen[2] (ychzhen@unicef.org)

[1] Social and Economic Policy Consultant, UNICEF Office of Research - Innocenti
[2] Social and Economic Policy Specialist, UNICEF Office of Research - Innocenti

Abstract: Measuring inequality of learning outcomes in a way that provides meaningful benchmarks for national policy while retaining a focus on those students who are 'hard to reach' and 'hard to teach' is a challenging but vital task in the light of the global post-2015 education agenda. Drawing on PISA 2012 data and its earlier rounds, this paper explores alternative approaches to measuring educational inequality at the 'bottom-end' of educational distribution within the cross-national context. Its main aim is to understand how far behind children are allowed to fall in their academic achievement compared to what is considered a standard performance in their country. Under the framework of relative (measured as achievement gap between the median and 10th percentile) and absolute (measured by the percentage of students achieving at a given benchmark) educational disadvantage it examines cross-country rankings as well as national trajectories with reference to overall academic progress. We find that on average across OECD countries around 11% of 15- year-olds lacked skills in solving basic reading, mathematical, as well as science, tasks in 2012, but variation across countries was large. The average achievement gap in mathematics across OECD countries between low-achieving and 'average' students stood at around 122 score points; in reading, at 131 score points; and in science, at 124 score points. This paper argues that understanding how the reduction in bottom-end inequality is achieved matters no less than the outcome itself, as it often reflects the level of support provided to low-achieving students. As our analysis shows, narrowing the achievement gap might be due to falling academic standards and have no direct benefit to the 'bottom group'.

Keywords: educational outcomes, low-achievement, measuring inequality, PISA.

Acknowledgements: the authors would like to thank their colleagues at the Social and Economic Policy Unit at the UNICEF office of Research, particularly Sudhanshu Handa and Emilia Toczydlowska for stimulating discussions and feedback at all stages of the Report Card 13 project. They are grateful to the members of the Report Card 13 Advisory board committee Francesca Borgonovi, Nevena Kulic, as well as Goran Holmqvist for the valuable suggestions. Special thanks to Dominic Richardson and John Jerrim who provided their insightful feedback and helpful comments on full drafts of the paper.

Country Abbreviations

Australia	AUS	Greece	GRC	New Zealand	NZL
Austria	AUT	Hungary	HUN	Norway	NOR
Belgium	BEL	Iceland	ISL	Poland	POL
Bulgaria	BGR	Ireland	IRL	Portugal	PRT
Canada	CAN	Israel	ISR	Romania	ROU
Chile	CHL	Italy	ITA	Slovak Republic	SVK
Croatia	HRV	Japan	JPN	Slovenia	SVN
Czech Republic	CZE	Republic of Korea	KOR	Spain	ESP
Denmark	DNK	Latvia	LVA	Sweden	SWE
Estonia	EST	Lithuania	LTU	Switzerland	SWZ
Finland	FIN	Luxembourg	LUX	Turkey	TUR
France	FRA	Mexico	MEX	United Kingdom	GBR
Germany	DEU	Netherlands	NLD	United States	USA

TABLE OF CONTENTS

Figures

Tables

1. INTRODUCTION

Two interlinked but distinctive concepts of educational inequality commonly take the centre stage of academic and political debate internationally: inequality of educational opportunity and inequality of educational outcomes. Both relate to unfair disparities between individuals or groups, but the priority placed on each can vary depending on educational policy agenda and prevailing political discourse. Focusing primarily on educational outcomes for 15-year-olds this paper is based on the normative premise that inequality among children is unfair since they are likely to have no control over resources and their circumstances. These two concepts are 'two sides of the same coin' (UNDP, 2013). Achieving success for all depends on providing a fair chance to the most disadvantaged through a 'level playing field'. In turn, poor educational outcomes compromise opportunities for the next generation of children.

First, equality of outcomes is important even if everyone has an equal starting point since we cannot ignore outcomes if they result in 'hardship' for some, even if opportunities were fair (Atkinson, 2015). In education this means that we cannot ignore the situation of children falling so far behind academically that it is deemed unacceptable on moral and child's rights grounds. Second, the unequal value of prizes in 'the race for success' and the process that determines their value matters. Some scholars argue however, that fair distribution of rewards might not happen through the market place because human capital is only partially tradable (Thomas, Wang & Fan 2001). Ensuring that educational outcomes are achieved in a fair way is vital for giving young people fair chances in the market place, which in turn contribute to social cohesion in a society. Third, some evidence from different disciplines demonstrates that gaps in educational achievement between rich and poor children emerge early in life (Feinstein 2003, Hansen and Joshi, 2007, Blanden and Machin, 2010, Waldfogel, 2013, Platt et al., 2014) with gaps between more able low socio-economic status and less able high socio-economic status (SES) converging or remaining stable over the child's life-course.[1] Intergenerational transfer of disadvantage through varied parental practices and investments (Katz et al., 2007, Lareau, 2003) means that inequality of educational outcomes today is likely to exacerbate inequality of educational opportunity for future cohorts of children.

This paper makes a contribution to the current debate and mainstream thinking about educational inequalities by focusing on 'bottom-end inequality' (Currie et al., 2010) – a concern about relative and absolute disadvantage of those at the very bottom of the academic distribution. Measuring 'bottom-end' inequality and educational disadvantage is important from a national as well as international policy perspective. Providing relevant evidence across a large group of industrialized countries helps to raise questions on explicit (further retraining, possible welfare support) and implicit (lost productive output) cost of educational disadvantage for the national economy. It also helps to raise national accountability against set educational goals and outcomes. In a global context of the post-2015 Education for All framework (UNESCO, 2012), it is becoming increasingly important to develop reliable indicators which reflect a focus on learning outcomes for the most disadvantaged. Moreover, the post-2015 human development agenda (SDGs) proposed a

[1] The issue of measurement error (so-called 'regression to the mean') has often been discussed in the context of these findings (see Jerrim & Vignoles, 2013).

major shift in development that would leave no one behind. This is driven by the normative principle that as countries progress towards academic excellence they should not allow their most academically disadvantaged be left behind as it will inevitably impede progress for all.

This paper builds the bridge from the UNICEF Report Card 9 (2010), which examined the situation of children 'left behind' in education, health, and material well-being relative to their peers in society. The research questions asked by Report Card 9 not only retain their relevance for the current global policy agenda in education but provide a timely stimulus for further discussion on how to measure and address educational disadvantage and inequality in a policy meaningful way. This paper examines:

- How far behind are children allowed to fall in their educational achievement compared to what is typical for their national context?
- What is the general dynamic between average academic performance and bottom-end inequality?
- What is the extent of academic disadvantage and how does it compare across 39 industrialized nations over time?

Using the OECD Programme for International Student Assessment (PISA) data sets for 2000, 2003, 2006, 2009 and 2012 PISA survey rounds, the analysis extends the scope of Report Card 9 (in the educational domain) to investigate the evolution of inequality of educational outcomes over time. Its purpose is to examine whether progress towards academic excellence can be achieved alongside low levels of bottom-end inequality and educational disadvantage. As a comparative study of national educational systems this paper provides evidence on, and contributes to, the discussion of the complex dynamics between equality of educational outcomes for children with the greatest educational needs and overall academic progress for all children.

We conceptualise inequality of educational outcomes as relative and absolute educational disadvantage. The former focuses on the disparity in academic achievement between 15-year-olds at the very bottom of national academic distribution and those with average academic performance. The latter identifies the severity and scale of 'bottom-end' educational disadvantage and helps to interpret the meaning of 'being left behind' through the level of skills and competencies acquired.

We find a relatively high score gap between median and bottom-end achievers across OECD countries, equivalent to around three academic years in any of the subjects. The evidence collected by this paper gives no ground to suggest an inevitable trade-off between average academic progress and bottom-end inequality. There was a positive correlation between median maths scores and the bottom-end achievement gap in 2012, but it was driven to some extent by a few country cases.

The ways in which countries reduce relative and absolute disadvantage do matter. Our analysis demonstrated various paths countries take as regards the level of support provided to the most academically disadvantaged children. Three central European countries (Estonia, Latvia, Poland) and Ireland showed the lowest relative and absolute disadvantage across the countries studied, demonstrating low levels of bottom-end achievement gap as well as a low share of children with cross-subject poor performance. At the other end of the spectrum, Bulgaria, Israel, the Slovak Republic and Sweden represent the educational systems with the highest levels of educational disadvantage under both indicators.

This paper is structured as follows: Section 2 underlines the conceptual framework and key definitions for our analysis. It also discusses various measures of inequality of educational outcomes used in the literature highlighting the rationale for the choice of the measures for this paper. Section 3 presents the data set and its limitations. Section 4 presents results for a) relative and b) absolute educational disadvantage. Section 5 brings those together in the combined country league tables and league matrix. Section 6 touches upon the relationships between macro-level context and our key indicators. We conclude with key messages and study limitations.

2. FRAMEWORK, DEFINITIONS AND MEASUREMENTS

This paper adopts the framework offered in the Innocenti Report Card 4 (UNICEF 2002), broadly operationalising inequality of educational outcomes through relative and absolute educational disadvantage. The Report Card assumes that reaching greater equality of outcomes through the support of the most academically disadvantaged children is an important policy task. The definition of relative educational disadvantage draws from Townsends' seminal work on poverty (1979: 15) in which he refered to relative deprivation as a lack of resources "so seriously below those commanded by the average individual or family that they are, in effect, excluded from ordinary living patterns, customs and activities". The conceptualisation of bottom-end educational inequality follows this relative perspective as it anchors the point of reference to the mid-point of academic distribution within a national context. It implies that children left too far below the 'national average' are relatively disadvantaged academically as they cannot command the same level of skills and competencies deemed standard in the society. This is different from other approaches which measure inequality of educational outcomes as the gap between the top and bottom educational performance or the gap between specific social groups (Duncan & Murnane 2011).

We consider education acquired by 15-year-olds to be a valuable set of skills and competencies which are then applied in everyday life, in familiar and unfamiliar contexts. In this sense, along with the disparity of learning outcomes between children at 'the bottom' and those in the middle, the bottom-end inequality of outcomes implies that some children fall below a certain minimum floor of academic standards. It can be argued that failure to reach this minimum (absolute) floor of literacy skills during formal schooling imposes risks on young people in different aspects of life, particularly in the labour market (McIntosh, S. 2004). We define this as absolute educational disadvantage which, along with the relative disadvantage, helps us to understand the situation of children at the 'bottom-end' of the academic distribution.

For the purpose of this study we define relative educational disadvantage as the achievement gap between the lowest and 'average' academic achievers. Absolute educational disadvantage is defined as falling below the internationally recognised threshold of low academic performance.

Below is the overview of specific measurement instruments for these two conceptual 'arms' of inequality of educational outcomes. We hypothesise that educational systems take various routes in their progress of reducing relative and absolute disadvantage, depending on specific policy choices rather than how well the system is performing overall: a) the support focused of the most academically disadvantaged children, b) inclusive progress with equal improvement among all low-performers, and progress while leaving the lowest performers behind.

The topic of educational inequality versus average academic progress at the country level has probably had the same fascination for educational research as income inequality versus economic growth explored in the economic development literature. The OECD Programme for International Student Assessment (PISA) publications emphasise that equality of outcomes and high average performance do not have to be mutually exclusive. The results of empirical testing of this premise are generally sensitive to the particular measures of inequality used (Hermann & Horn 2011).

Using years of schooling as an indication of educational attainment and an education Gini index, Thomas, Wang & Fan (2001) find that inequality falls as average years of education increase. Using standard deviation as the measure of inequality, the same study finds a U-shaped relationship with inequality: first rising with average years of schooling and then falling. With a very low turning point of 6-7 years of schooling, the authors claim that the majority of industrialized countries would 'pass' this turning point by the end of secondary schooling or at the higher levels of educational achievement.

Freeman, Machin & Viarengo (2010), using 1999 and 2007 Trends in International Mathematics and Science Study (TIMSS) data, find a negative association between achievement and inequality: higher overall achievement is associated with lower inequality in test scores across countries. The authors refer to this relationship as 'virtuous' equity-efficiency trade-off in improving educational outcomes. Micklewright & Schnepf (2006) use a gap between 95th and 5th percentile transformed into z-scores in six tests in PISA, TIMSS and PIRLS to investigate the same relationship. They find what they call a 'reasonably clear pattern' where within-country differences are highest where average achievement is lowest. Their sample was limited to 21 industrialized countries excluding Chile, Mexico and Turkey, covered by the present study. None of the above-mentioned papers focused specifically on 'bottom-end inequality' as defined by this paper.

Although it cannot be denied that some inequality of educational outcomes is not only a reflection of students' natural abilities, but also a pre-condition for student's motivation, self-efficacy and a necessary stimulus for academic success (Brighouse, 2014), educational systems seem to differ in the way they nurture their top performers and low achievers. This is likely to have a direct impact on the distribution of academic scores and overall academic progress.

2.1 Measuring relative educational disadvantage

The item response model and standardisation procedures applied in PISA data have a profound effect on the selection of the measure of inequality of educational outcomes. Three different approaches have been used so far in the literature. We empirically test some of the approaches and compare their consistency in country rankings and error terms. The final choice of the measure has to be taken based on the corresponding definition and analytical goals of this paper.

Educational achievement data are recorded on a continuous scale. This motivated researchers to apply inequality measures such as ratios and indices developed for the analysis of income inequality. For instance, using the Trends in International Mathematics and Science Study (TIMSS) of 1999 and 2007, Freeman, Machin and Viarengo (2010) used the ratio of the difference between a country's 95th percentile and 5th percentile scores to the median as a measure of dispersion. Based on this measure (the rationale for which is not explicitly discussed) the authors find that lower inequality in test scores was associated with higher average achievement scores across countries. The UNICEF Report Card 9 (UNICEF, 2010), which examined bottom-end inequality in material well-being, education and health, also used the relative ratio (50th-10th/50th *100) in analogy with the income inequality measure.

Some researchers warned against adapting standard measures of income inequality to data from educational surveys such as PISA. Micklewright & Schnepf (2006) noted that as achievement scores are estimates from an item response model, rather than children's actual survey responses, so their nature is different from that of income data. The difference is that income data are observed directly, reflecting factual distribution in the population, while achievement scores have an imposed normalised distribution. Therefore, Micklewright & Schnepf argue that a distribution of test scores cannot be interpreted as representing the distribution of skills or knowledge in the population of 15-year-olds.

Ferreira & Gignoux (2011) extended this argument pointing to the problems associated with further transformations of the achievement score through standardization procedures. Notably, the standardisation results in both the translation of the mean and the rescaling of the dispersion. This violates the assumption of invariance of the distribution before and after the transformation. They further argued that one cannot say with confidence that the inequality measure, which is defined over the standardised test score, will be the same as the one defined over the original distribution. Ferreira and Gignoux propose variance or the standard deviation (SD) as an alternative measure that does not have the above-mentioned problems. This approach has also been used recently by other studies using PISA data (Choi & Jerrim, 2015; Blanden & McNally 2014).

Although methodologically the SD might be the most appropriate measure of variance in educational outcomes cross-nationally, it is not an easy concept to interprete and communicate to policymakers or education system professionals. Moreover, by definion it describes the variance or dispersion of scores around the mean, not median. Even though the mean and median will produce very similar data due to the distribution being normalised, the median or scores in the 'middle' is a reference point for relative educational disadvantage based on its definition.

Another approach taken in the literature is the test-score difference between the students represented at different test score percentiles. The PISA 2009 report on inequality (OECD 2010) asserts that the score point difference across percentiles is a valid measure of performance gaps within countries. Moreover, the same report refers to the gap between the median and the 10th percentile[2] as the measure of the achievement gap at the bottom of the distribution. Micklewright and Schnepf (2006), using TIMSS and PISA 2000 round data, examined within-country differences in inequality of learning outcomes across 21 countries using absolute gaps between percentiles (P95-P5). Report Card 4 (UNICEF, 2002) uses the gap between the median test scores[3] and the 5th percentile to report the 'bottom end inequality' as educational disadvantage.

We have empirically tested both alternative measures: the standard deviation (SD) from the mean proposed by Ferreira & Gignoux (2011) and the percentile range (50-10 percentiles). In addition, we tested a coefficient of variation (SD/mean). The results are presented in Table A.1. (Appendix). Not surprisingly, given the normal distribution of educational data imposed through assumption of the item response model, standard deviation and the percentile gap 50-10 produce very similar country rankings with consistent groups of top and bottom performers. An exception is mathematics in the Netherlands, which as we will see later has a distribution somewhat skewed to the low-end, thus a lower ranking on the 50-10 percentiles gap measure. Meanwhile, the achievement gap measure produces higher statistical error terms than the SD measure.

The country ranking based on the coefficient of variance (SD/mean) is not consistent with a ranking based on SD or percentile range. The difference is primarily observed among high performing countries which generally do better under the former. For example, Korea is ranked 34th and 35th under SD and the gap measures, but is ranked 11th under SD/mean. The PISA 2012 report (OECD, 2014a) indicates that Korea achieved recent progress in mathematics by focusing on top performers rather than low achieving students. Therefore, it appears that SD/mean produces a misleading picture for our analytical task of focusing on bottom-end inequality. Moreover, similarly to the ratio of 50-10th percentile gap to the median, which we tested separately, it is likely to suffer from all the problems of ratios used as a measure of inequality identified by Ferreira & Gignoux (2011).

We conclude that within our conceptual framework and with the chosen definition of bottom-end inequality as the relative position of low-achievers compared to what is considered typical in the national context, the percentile range (P50-P10) of the relative achievement gap is the appropriate measure of bottom-end inequality or relative disadvantage. But its interpretation will be within the context of progress in average achievement and the characteristics of absolute educational disadvantage.

2.2 Measuring absolute educational disadvantage

Defining a measure of absolute educational disadvantage is always an arbitrary task. Educational achievement as an outcome is embedded in the national context. It varies across national educational systems depending on the standards, scales, tests and measurement methods applied. Therefore it is difficult, if not impossible, to make cross-country comparison based on achievement levels defined within a national context.

[2] The 10th percentile is the score reached by 9 out of 10 students, but not reached by the lowest performing 10% (OECD, 2010).
[3] The 50th percentile, also known as the median score, is defined as the score that half of the students in the country do not reach and the other half exceed. The median student is the student in the middle of the performance distribution for each system.

PISA developed six proficiency levels defined for each subject for international comparative analysis based on a common scale. Each level reflects 'aspects' of each subject defined independently by experts in the field. For example, for reading, this includes accessing and retrieving information, integrating and interpreting texts, and reflection and evaluation. This provides a benchmark of absolute advantage or disadvantage in terms of functional literacy rather than the formal levels of education. This can generally be considered as a proxy of a country's future productivity and competitiveness (UNICEF Report Card 4). Proficiency levels were used, for example, to develop key indicators of learning outcomes in the post-2015 agenda of the Education for All Steering Committee (EFA SC) recommendations (Technical Advisory Group EFA, 2014).

Proficiency levels were designed with continuity in mind. They do not have natural breaking points to mark borderlines between stages along the scale (OECD, 2014b). Acknowledging that proficiency levels are simply a matter of degree, PISA defines stages or levels of learning outcomes as milestones of achievement, which are easy to communicate to, and be understood by, students, parents and educational professionals. One proficiency level roughly corresponds to 73 score points (OECD, 2010 Volume II: 27).

PISA defines low academic performance as acquiring scores below the threshold of proficiency level 2 in each subject. This corresponds to fewer than 420 score points in mathematics, 407 score points in reading and under 409 points in science. Based on PISA 2012 data, on average across OECD countries, less than a quarter of all children in the sample fall below the level 2 threshold: 23% of children do not reach level 2 in mathematics, 18% in reading and 18% in science. PISA reports emphasise that this is a particularly vulnerable group, at risk of early drop-out and non-progression to post-secondary or tertiary education (OECD, 2010).

Given our focus on 'bottom-end' inequality with the reference to the 10th percentile group, we have examined the characteristics of more severe deficiency in skills and competencies. PISA defines the threshold of 'below level 1' in mathematics as scores below 357.7, in reading less than 334.7, and in science 334.9. On average across OECD countries, 8% of children in mathematics, 6% of children in reading and 5% in science fall under this threshold.

Using 'below proficiency level 2 and level 1' thresholds we structure the analysis from two angles:

a) 'Depth' or severity of absolute educational disadvantage is defined as the concentration of severe educational need. It is measured as a proportion of children performing below proficiency level 1 relative to the whole group of low performers (below proficiency level 2).

b) 'Breadth' of absolute educational disadvantage is defined as overlapping, cross-subject low performance under proficiency level 2 threshold. It is measured as a per cent of children below level 2 in all three subjects.

We find that the 'breadth' or cross-subject low performance measure is a robust measure producing country rankings consistent with the ranking based on the PISA formal definition of low performance. But for the purpose of this paper it helps to distinguish a group of children and young people closer to the '10th percentile' defined as our 'bottom-end group'. Along with the bottom-end achievement gap it was therefore chosen to be the base for the combined country league tables.

3. DATA AND METHODS

3.1 PISA data and its limitations

This paper uses the OECD Programme for International Student Assessment (PISA) data sets for 2000, 2003, 2006, 2009, 2012 survey rounds. Although other international datasets such as TIMSS (the Trends in International Mathematics and Science Study) and PIRLS (the Progress in International Reading Literacy Study) are also suitable for the analysis, the choice of PISA data is stipulated by a wider country coverage and the analytical starting point – to replicate to the extent possible the analysis presented in Report Card 9, which was based on PISA data.

In total, 39 countries are included in the analysis (OECD members and partner countries). Countries were selected based on data availability and the inclusive criteria of a wide geographic representation. The country coverage of UNICEF Report Card 12 (2014) was taken as a reference point. The individual and combined country rankings are based on the PISA 2012 survey data and include 37 countries. Mexico and Turkey are excluded from the final league table rankings due to high rates of children not completing upper-secondary education[4] and a possibility of system bias. However, they are included in all descriptive analysis that precedes the league tables. Country coverage for trend analysis is discussed in detail below.

The PISA survey employs a two-stage stratified sample design; the units in the first stage sampling are individual schools.[5] The total sample of the PISA 2012 round consists of 458,491 children aged from 15 years 3 months to 16 years 2 months. Consistent with the PISA approach we use '15-year-olds' as shorthand for the targeted population represented by the survey. According to the official PISA definition, the sample of children includes those in grade 7 or above within corresponding national educational systems. In PISA 2012, 32 countries administered paper-based as well as computer versions of mathematics assessments. Although there can be differences in results between the two types of assessment (Jerrim, 2016), this paper makes an implicit assumption of homogeneity of assessment instruments. The country which showed the most marked difference in PISA score points (Shanghai-China) is not included in our analysis.

PISA collects data on students' competencies in mathematics, reading and science. It aims to measure broad skills and the ability to apply them in real life situations. So the scores reflect skills and competencies, rather than knowledge of particular facts, within an internationally recognised curriculum. Competencies are modelled as latent variables using a one-parameter Rasch model (item response model). [6]The unobserved literacy scores are 'unbounded', i.e. having nether top/maximum nor bottom/minimum scores. The average competence score is standardised (at the year when a subject first became a major domain) to produce the OECD mean of 500 and standard deviation of 100 score points. Thus the dataset is specifically designed for international comparative analysis.

[4] For statistics on upper secondary completion (graduation rates) across the examined countries see OECD (2014). Education at a Glance 2014: OECD Indicators. Table A2.1a.
[5] For details on sampling procedure see PISA 2012 Technical report (OECD, 2014b).
[6] For more information on PISA procedures for scaling cognitive data see PISA 2012 Technical Report. Chapter 9, available at http://www.oecd.org/pisa/pisaproducts/PISA%202012%20Technical%20Report_Chapter%209.pdf

This paper recognises that PISA data was designed and implemented for a very general overview of students' achievements across many countries rather than to give detailed and encompassing insight into the quality of schooling in each country. But its focus on the adolescent age group, wide country coverage, availability of data across many time points and consistency of methodology makes it the most suitable data source for this analysis.

Some specific limitations of PISA data have direct implications for our analysis. The PISA survey is designed to be representative of the 15-year-old student population attending schools. But the drop-out rate by the age of 15 can be substantial in some countries. For example, in 2012, the enrolment of 15-19-year-olds in education in Turkey was only 59%, and in Mexico was 56% (OECD, 2013). This can cause system-wide bias (Richardson & Ali 2014) or underrepresentation of the most disadvantaged children in the sample. Moreover, according to the new definition of the student population introduced by PISA organisers in 2012 (OECD, 2014b), children of the targeted age who are enrolled in a grade lower than the grade 7 threshold due to repetition or for other reasons, are not included. This can also affect estimates for the group of lower achievers. Reported coverage rates (the ratio of the population of 15-year-olds) can be low. To confront this issue, the paper excludes countries with reported high drop-out rates (namely Turkey and Mexico) from the final combined league tables.

Student achievement data in the PISA dataset contains five plausible values for each student in each subject. These are multiple imputations of the unobserved latent achievement for each student. They are estimated under assumption of a flat linear regression with all students' background characteristics entered as predictors (Monseur & Adams, 2009). PISA datasets also provide replicate weights appropriate for this particular survey sample design. Using replicate weights ensures the analysis produces accurate and unbiased estimates of standard errors. All analysis in this paper has been done using the 'repest' package for STATA, with commands specifically for use with complex datasets like PISA, using replicate weights and plausible values.

3.2 Trend analysis

A simple requirement for any trend analysis is the consistency of measurement instruments and the characteristics of outcome, which we are trying to measure. There are a number of issues related to the comparability of PISA data across cycles. For instance, PISA survey design changes the subject coverage from cycle to cycle. The domains of reading, mathematics and science rotate as a major focus every three years. The scales for mathematics, reading and science are first established when a subject is a major domain. The scale is set to have a mean of 500 and standard deviation of 100 (for the pooled, equally weighted OECD countries). Following a simple requirement of using a consistent scale for each subject, trends can only be reported from the round when the scale was established. Reading was a major domain in 2000, mathematics in 2003, and science in 2006. Our key indicators are comparable across major and minor domains.

The second issue relates to harmonisation of country coverage across rounds. PISA organisers specify that a number of national data have to be excluded from the trend analysis due to countries' failure to meet PISA sample response rate requirements, different issues of test administration procedures or other technical issues. This is the case for reading performance in the year 2000 for the Netherlands and Luxemburg,[7] for the United States in reading in 2006[8] and the United Kingdom for 2000 and 2003.[9] For these countries the trend analysis in reading was adjusted accordingly. Following OECD recommendations, Austria is excluded from all trend analysis due to concerns over a data collection issue in 2009.

The inclusive approach taken by this study in terms of country coverage also means that there is a great variation of available national data depending on the year of a country joining PISA. Thus, the Slovak Republic and Turkey joined PISA in 2003; Bulgaria, Chile, Croatia, Estonia, Israel, Lithuania, Romania and Slovenia joined the PISA survey in 2006. Thus data for these countries are not available for earlier rounds posing limitations on the time period for a cross-country comparison, particularly for maths and reading which are meant to be examined from 2003-2012 and 2000-2012 respectively. We use the earliest data points available (providing their validity) for presenting individual country rankings. However, for consistency of representation of relative and absolute disadvantage across 38 countries (excluding Austria), although other reference points are available we focus on the period between 2006 and 2012 due to availability of data for all countries over the last three PISA rounds.

This paper analyses trends in relative and absolute educational disadvantage as an absolute difference in our key indicators between the observed periods. This simplest of approaches applied by PISA (OECD, 2014b) is motivated by the transparency of absolute change we can observe between each cycle. This approach seems appropriate given our focus on the last three rounds (due to a wide country coverage) and complementary representation of selected country trajectories from earlier available points.

Consistent with the PISA approach, an OECD unweighted average is used as the focus of the paper to compare performance across educational systems. An OECD average is calculated for all key indicators. It corresponds to the arithmetic mean of the respective estimates of OECD countries included in the sample. Standard errors of the OECD average are calculated from the average of the variance across the measurements, assuming pooled variance.

[7] The Netherlands did not meet the response rate requirements in PISA 2000, while in Luxemburg testing conditions in 2000 were different from those in the following years (OECD 2005).

[8] No reading results are available for USA PISA 2006 due to problems with test administration.

[9] For the United Kingdom, the PISA 2000 and 2003 samples did not meet the response rates standards. After collecting some additional evidence, PISA reported UK results in 2000 but not in 2003. But, after the introduction of a more strict response rate standard, both rounds had to be removed from the trend analysis. More concerns have been raised recently about the reliability of later rounds of PISA data for the UK due to the administration of test procedures and non-response (Jerrim 2013). We include later rounds for the UK in our analysis but recommend caution in interpretation of results.

4. RESULTS

Students' test scores in mathematics, reading and science are highly correlated.[10] Not surprisingly, the correlation between maths and science is stronger than either of these subjects with reading. Therefore, due to space limitation, we chose to present results by focusing either on mathematical or reading performance whenever the substantive differences between them are pronounced. Complementary statistics for other subjects can be found in the Appendix.

4.1 Relative educational disadvantage

4.1.1 Academic distribution

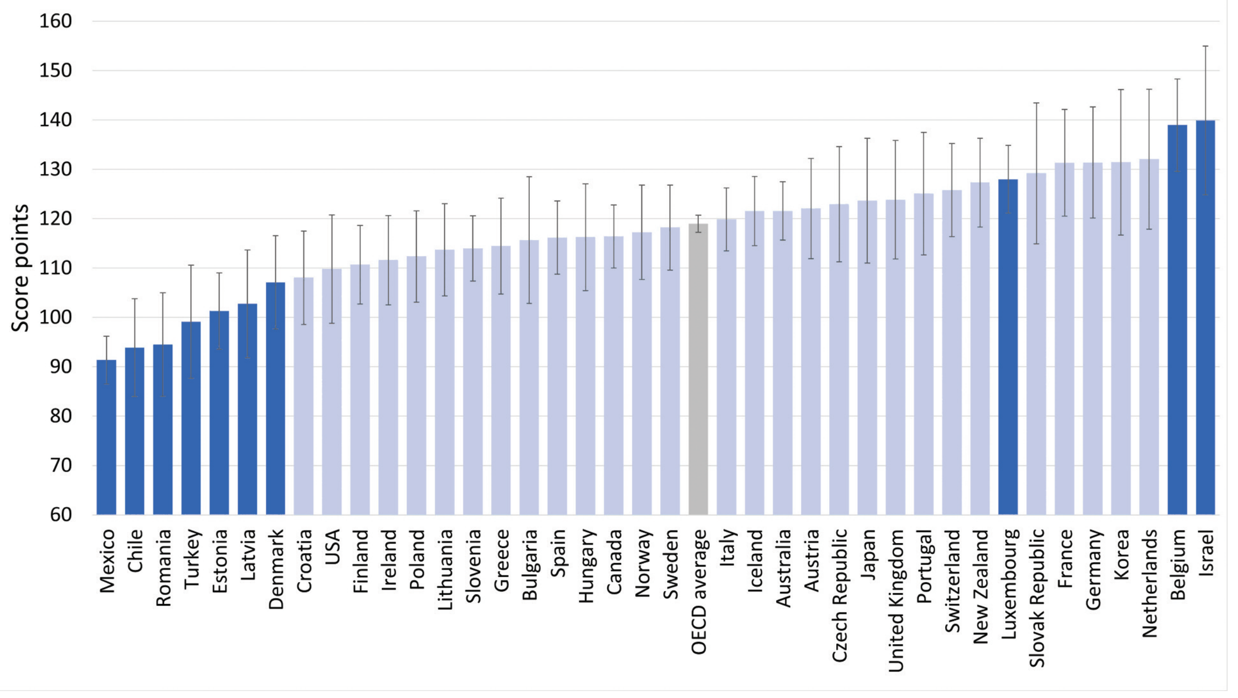
Our first step is to examine the spread in the test scores around the median, with the focus on the dispersion at the low end of the academic distribution. Figure 1 shows bottom-end achievement gap between the median and the 10th percentile in mathematics across 39 countries in 2012. The average absolute gap across the OECD countries stood at around 119 score points. Countries are presented in three groups according to their position in relation to the OECD unweighted average. We find that Chile, Denmark, Estonia, Latvia, Mexico, Romania and Turkey have relative disadvantage levels significantly below the OECD average. Meanwhile, for Belgium, Israel and Luxemburg the achievement gap between the median and the 10th percentile is significantly higher than the OECD average. There are also many countries in 'the middle' of the graph with the bottom-end achievement gap not significantly different from the OECD average.

To fully appreciate the scale of variance between the average and bottom performers in mathematics within a national educational system as well as cross-nationally, it is helpful to translate the score point difference into the years of schooling (Willms 2001). PISA reports specify that 41 scores correspond to the equivalent of appropriately one year of formal schooling (see Annex A1, OECD 2014b). Israel is the country with the largest achievement gap in mathematics of 140 score points, followed very closely by Belgium (Figure 1 page 20). Israel's achievement gap between a student at the median and bottom 10th percentile is the equivalent of nearly three and a half years of schooling. Even in Chile, Mexico and Romania – the countries with the smallest low-end percentile gap of 91 and 94 score points, it still accounts for over two years of schooling. This is noticeably lower than the average for the OECD countries (almost three years).

France, Germany, Korea, the Netherlands and Slovak Republic are very close in ranking position with the bottom-end achievement gap in mathematics close to the OECD average. Despite their proximity in the ranking, these countries have marked differences in the way their educational system nurtures its top performers and supports low achievers. This can be seen better through the shape of the achievement score distribution for all students: the dispersion of mathematical performance at the top (90th percentile) and bottom tails of the distribution (10th percentile) relative to the national median (Figure 2 page 21).

[10] Mathematics and science r: 0.93, p<0.001; reading and science: r: 0.93, p<0.001; reading and mathematics r: 0.90, p<0.001.

Figure 1 – ***Score point achievement gap between children at the 10th percentile and children at the median (50th percentile) in mathematics in 2012***

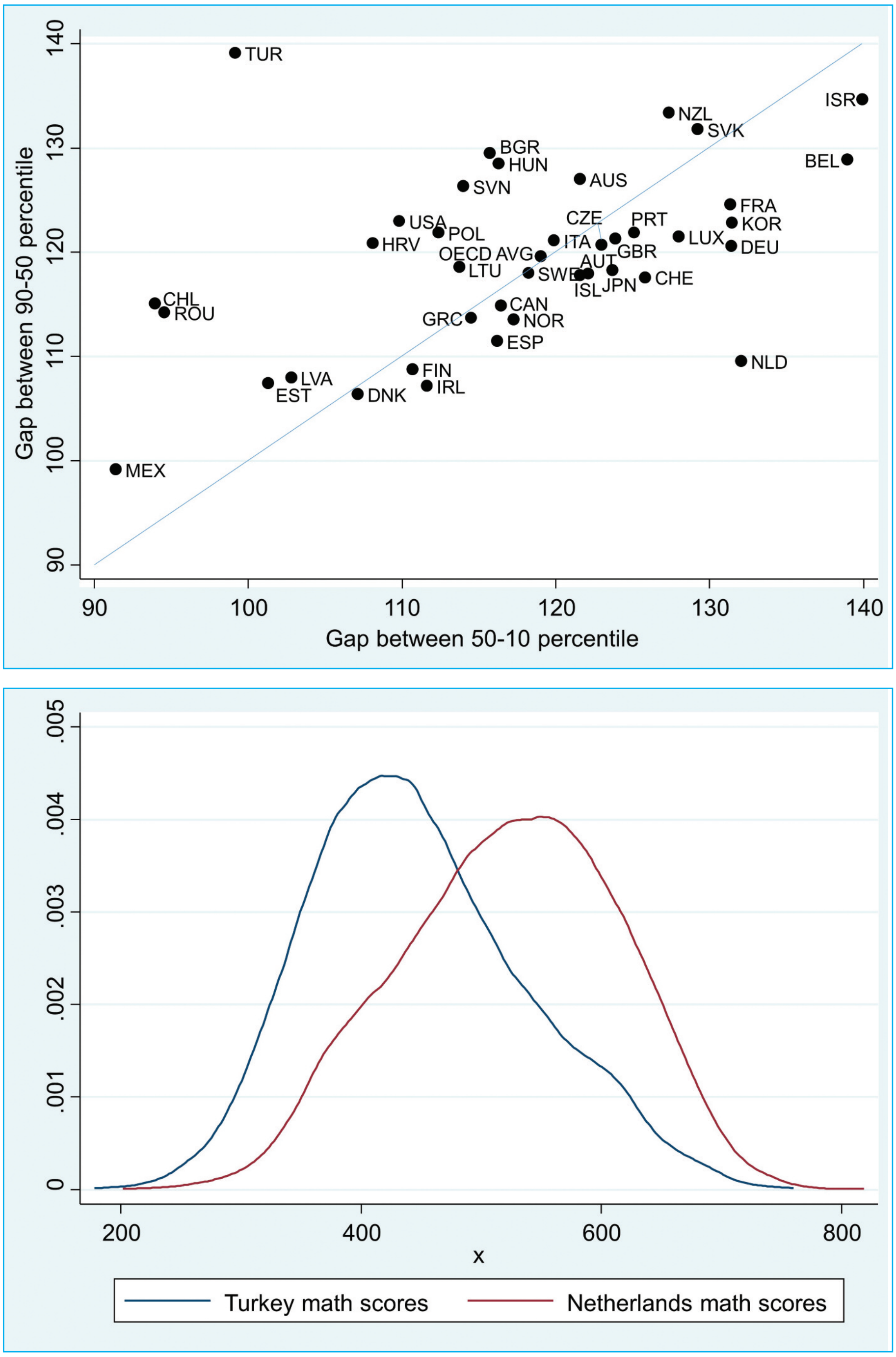


Note: OECD unweighted average. Bars marked in dark blue show a statistically different result from the OECD average.
Source: PISA 2012.

The distribution of academic achievement is typically normalised (due to the standardisation procedures). Despite this, we find a number of countries with very different score variations among top and low achievers. For example, the Netherlands has disproportionally large inequality at the bottom of the distribution (132 score points) compared to countries with a similar spread between the 90th and 50th percentile (Denmark, Estonia, Finland, Ireland and Latvia). This means that the Netherlands' educational system permits its academically weakest students to fall further behind in mathematics by around one extra academic year compared to a similar cohort of low-achieving students in Estonia or Denmark.

In contrast, Turkey has a very tight spread at the low end but one of the largest variations at the top despite the low levels of average achievement (Figure 2). Turkey's low end achievement gap of 99 score points, or two and a half years of schooling, falls far below the OECD average. However, this needs to be interpreted with caution. The country's high dropout rates by the age of 15 (OECD, 2013) can result in selection biases associated with enrolment and repetition rates (Ferreira and Gignoux, 2011). This could have an impact with the measure underestimating the spread of scores at the bottom of the distribution.

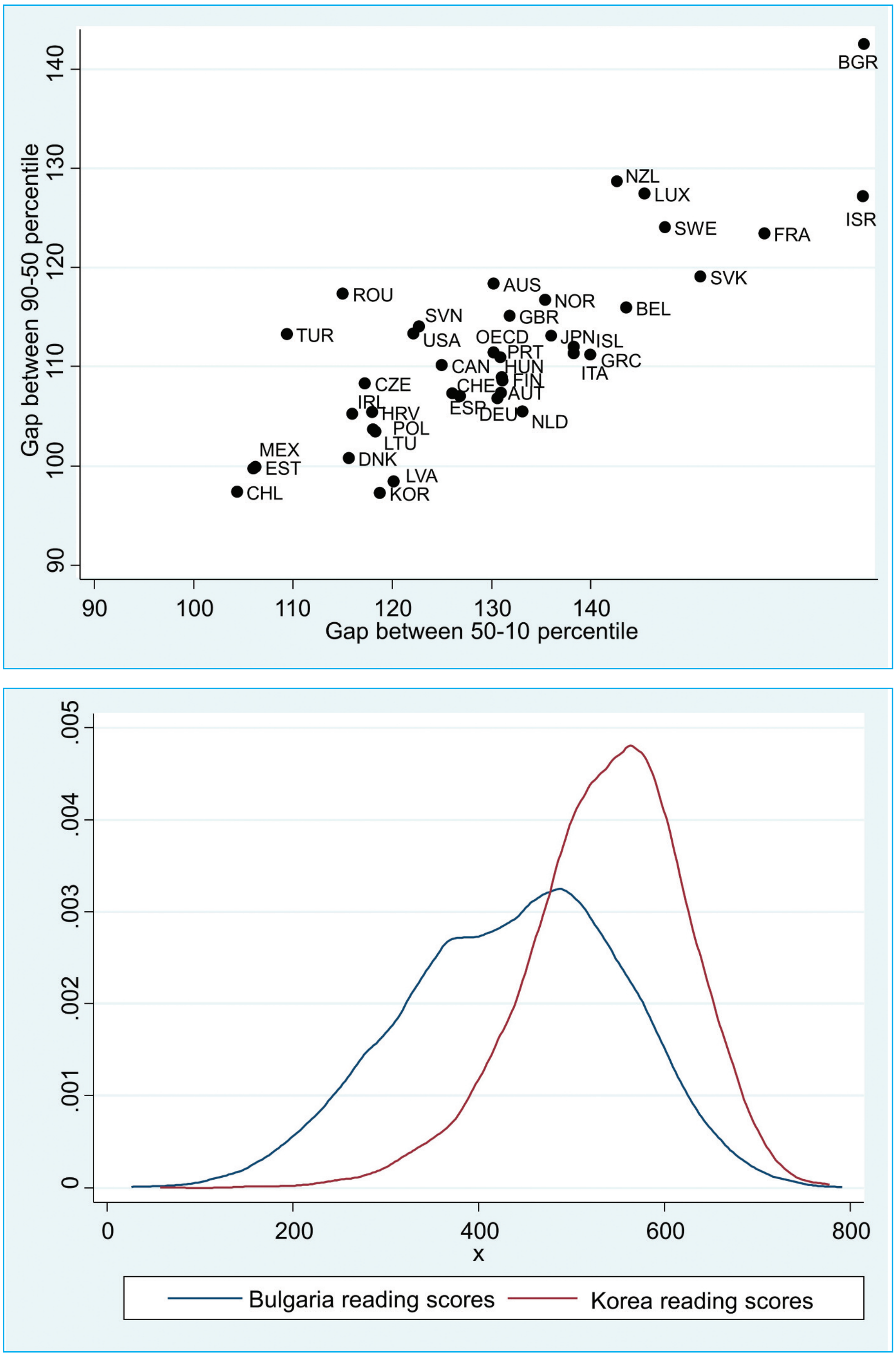
Figure 2 – ***Variation of test scores in mathematics: 90-50 percentile versus 50-10 percentile gaps***



Note: Mathematics: r= 0.93, p<0.001; science r=.93, p<0.001. *Source:* PISA 2012.

In reading, the distribution of academic performance is somewhat different to mathematics (Figure 3 page 22). The majority of countries have a wider achievement gap between the 50th and 10th percentiles than between the 90th and 50th. Bulgaria is an exception with a wide dispersion of scores on both ends around the median. The achievement gap between Bulgaria's median test scores and those at the bottom in reading is 167 score points. This is the equivalent of four years of schooling, one year more than the average in OECD countries (131 score point gap). The gap between Bulgarian top and median performers in reading is smaller. At 142 score points it is equivalent to 3.4 years of schooling.

Figure 3 - ***Variation of test scores in reading***



Note: Reading r=0.95, p<0.001.
Source: PISA 2012.

Korea, on the other hand, has a tight distribution of scores between the 50th and 10th percentiles in reading but an even smaller variation at the top. Moreover, the Kernel density plot in Figure 3 shows that despite very tight distribution, children at the bottom-end have reading scores relatively distant from those in the middle. The shape of the distribution at the bottom in Bulgaria and Korea is also different indicating that the handful of low achievers in Korea might feel more marginalised and /or pressurised than their peers in Bulgaria since their reading scores are so much lower compared to those in the middle. This also can possibly be linked to high pressure in school and at home to achieve high marks. We will further examine some characteristics of low academic performance in Korea and other countries in the section devoted to the absolute educational disadvantage.

Overall, we find that too many nations allowed their low achievers to fall too far from what is considered a norm of academic achievement in the society. In reading, it can be as high as four full years of schooling, as in Bulgaria and Israel. On average across OECD countries the gap between the average achievers and those at the 'bottom' equates to a staggering 3 years of schooling in reading according to PISA methodology. This is similar in mathematics and science. Even the countries with the lowest relative disadvantage (Chile, Estonia and Mexico) have achievement gaps equivalent to two and a half years of schooling. But to understand the full scale and meaning of relative educational disadvantage in the comparative cross-country context, we have to relate the achievement gap to the average levels of countries' academic achievement.

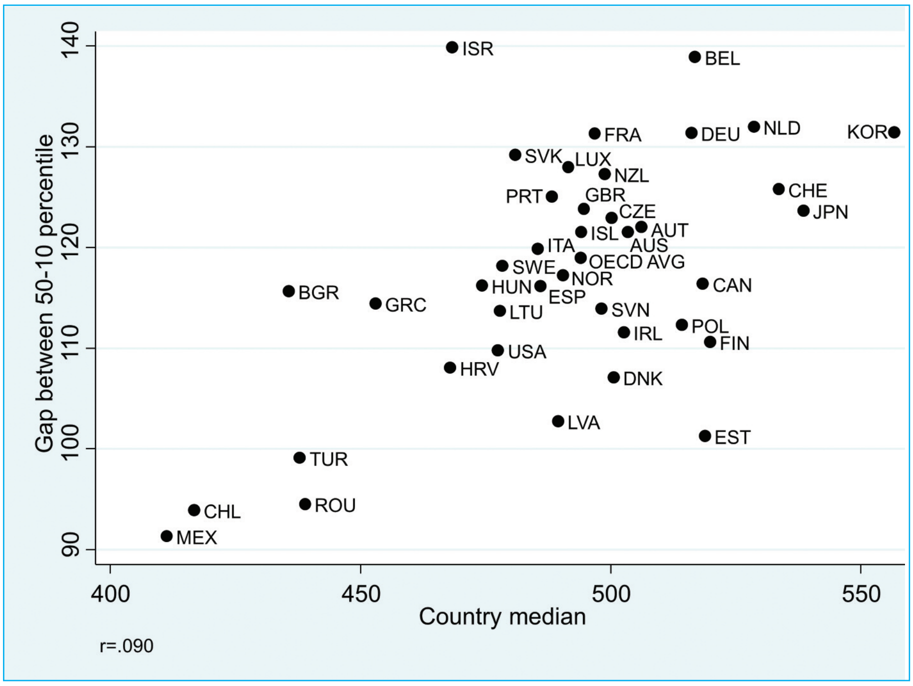
4.1.2 Relationship with the median

A first look at the relationship between the relative educational disadvantage and the country's median performance in mathematics (Figure 4 page 23) reveals statistically significant positive correlation coefficient (r=0.57), which is even weaker in science (r=0.37) but no longer significant in reading. Spearman's correlation coefficient was smaller in maths (r.=0.44), science (r=29), while no correlation was found in reading. It appears that countries with higher median performance in mathematics and science have somewhat higher bottom-end inequality. To some extent, the association is driven by countries like Chile, Mexico and Romania with very low scores, and also by the very small variation at the bottom end of the distribution. Indeed, excluding these countries makes the association in science not significant and reduces the magnitude in maths ($r=0.33$, $p<0.05$).

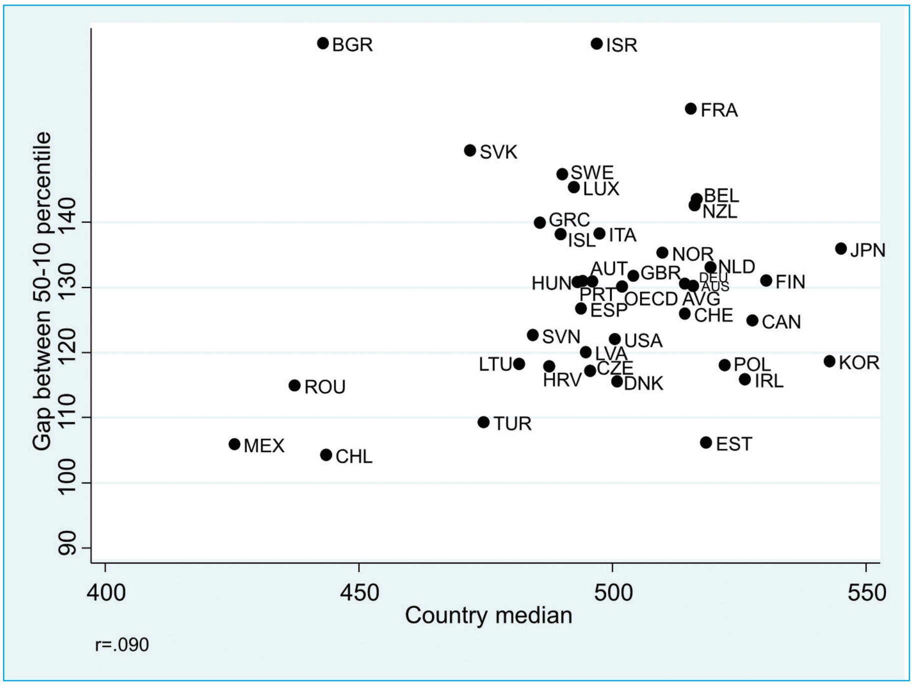
A number of countries with similar median scores have very different achievement gaps at the bottom. For instance, we find Bulgaria has a median mathematics score similar to Romania and Turkey (435 compared to 438 and 437 respectively) but a much higher bottom-end achievement gap: 115 score points in Bulgaria, 94 in Romania and 98 in Turkey. On the other end of the spectrum is Estonia. Being one of the top performers in terms of median scores in maths, Estonia's bottom-end inequality is one of the smallest (101 score points). The source of educational inequalities in Estonia is attributed primarily to the disparities found between non-native speakers (primarily Russian-speaking) and ethnic groups of children (Lindemann, 2013). Narrowing the gap between rural and urban schools through more equitable distribution of funds (OECD, 2014a), increasing the quality of teaching in the Russian-medium schools through monetary incentives to teachers are among some of the measures implemented in recent years to address achievement gaps. As a result, according to the Estonian Ministry of Education and Research (2015), achievement in Russian-medium schools improved twice as much as the Estonian-medium schools between recent PISA rounds.

Figure 4 – ***Relationship between median scores in a) mathematics, b) reading and 50th -10th gap***

A. Mathematics scores



B. Reading scores



Source: PISA 2012

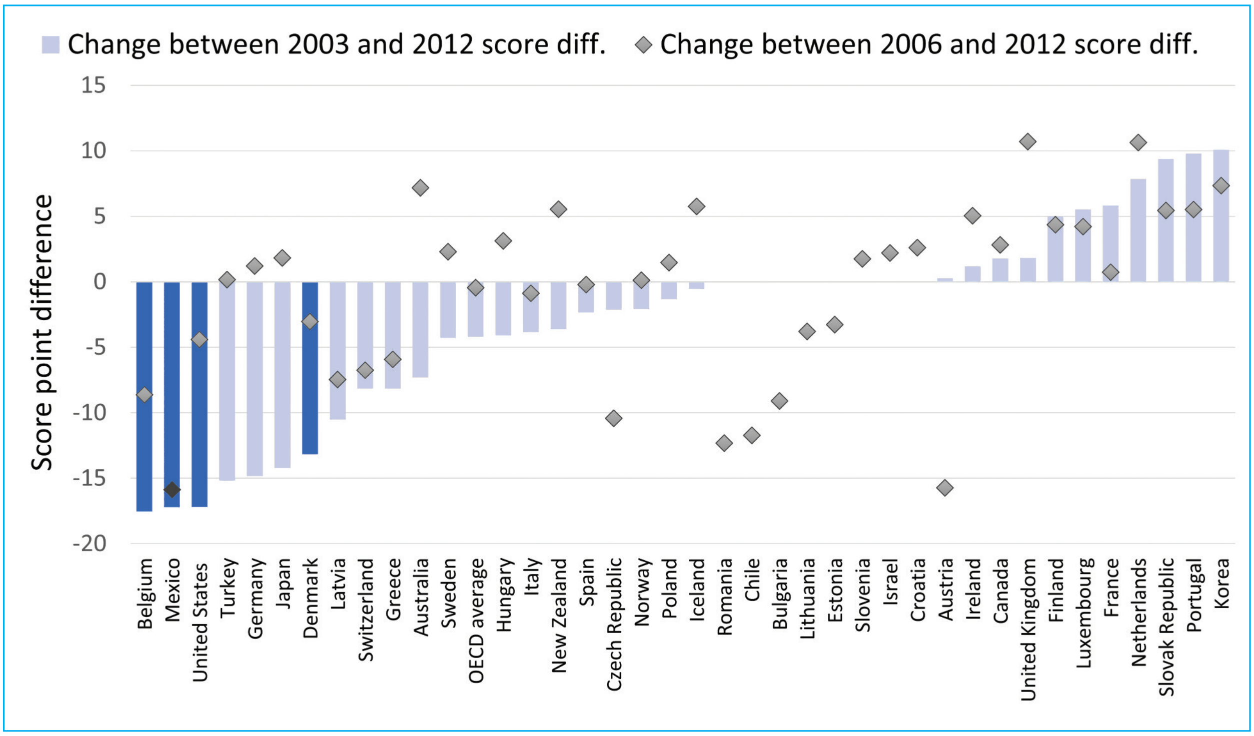
The lack of strong association between average academic performance and the bottom-end inequality (50th-10th achievement gap) can be taken with some cautious optimism. This suggests that there is no ground for us to assert an inevitable trade-off between higher national standards of achievement and leaving some students behind. Indeed, as the educational systems of Denmark, Estonia and Latvia have shown, it is possible to combine strong average academic performance with a small achievement gap at the bottom of the distribution.

4.1.3 Diverging trajectories of relative educational disadvantage over time

The change over time in achievement gap between students at the 10th percentile and those at the median indicates the extent to which the educational system is concerned with relative educational disadvantage and its lowest achievers. For many countries the change was non-linear between PISA rounds and rather mixed, more so for maths and science than for reading.

Among 38 countries included in the trend analysis[11] 20 countries have a downward trend in relative educational disadvantage in mathematical literacy between 2003 and 2012. On average across OECD countries the achievement gap was reduced by 4.2 score points. The largest contraction in the gap (above 10 score points) is observed in Belgium, Denmark, Germany, Japan, Latvia, Mexico, Turkey and the United States, but only in Belgium, Mexico and United States is the change statistically significant. 10 countries widened the achievement gap over the same period but the scale of the change is smaller than among the above-mentioned group. The largest increase between 2003 and 2012 by 10 score points is observed in Korea. It is followed by Portugal and the Slovak Republic with increases in the gap between 9 and 10 points. Of eight countries for which data is only available from 2006, five reduced the gap with the highest reduction seen in Romania (12 points), Chile (12 points), and Bulgaria (9 points).

Figure 5 – ***Change in the achievement gap between children at the median and the 10th percentile in mathematics***

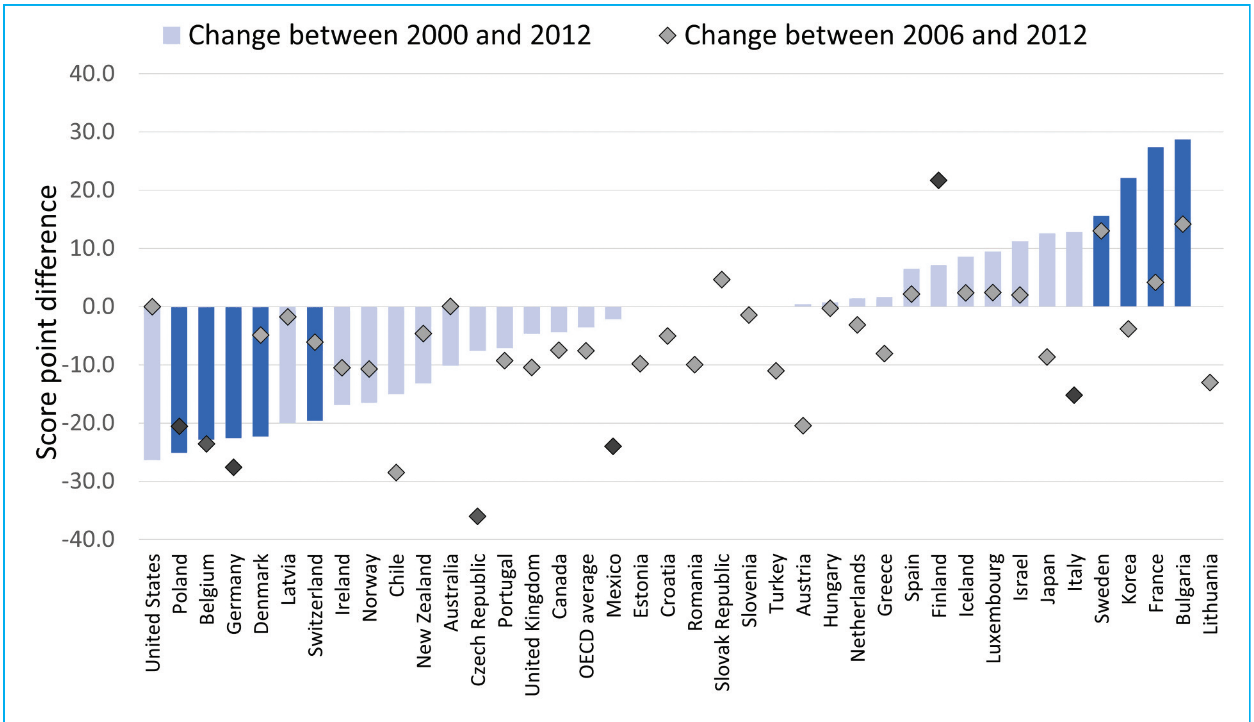


Note: OECD unweighted average. Statistically significant values are marked in a darker colour.
Source: PISA 2003/2006/2012

[11] Austria is excluded on the grounds discussed in the section 'Data and Methods'.

In reading, eighteen countries out of 32 for which data is available have a positive trend in relative educational disadvantage between 2000 and 2012. Figure 6 shows that the largest and statistically significant contraction in relative disadvantage in reading was achieved by six countries: Poland 25 (score points), Belgium and Germany 23 (score points), Denmark 22 (score points) and Switzerland 20 (score points). Over the same period, fifteen countries showed a trend in widening the bottom-end achievement gap. Statistically significant increase is observed in Bulgaria 29 (score points), France 28 (score points) and Korea 22 (score points)They are followed by Sweden (16 score points).The Slovak Republic increased the gap by 26 score points between 2003 and 2012 (not presented here) and Bulgaria by 14 points between 2006 and 2012.

Figure 6 – ***Change in the achievement gap between children at the median and the 10th percentile in reading***



Note: OECD unweighted average. Statistically significant values are marked in a darker colour.
Source: PISA 2000, 2006, 2012 data.

Overall, it appears that the majority of countries made progress in reducing relative educational disadvantage over time. But the reduction in the achievement gap between children at the median and at the 10th percentile could happen in many different ways. Children at the bottom can be 'lifted up' so that their achievement level is closer to the national average. But it is equally possible, at least hypothetically, for the national average to fall with the low-achievers' score remaining constant.The bottom-end group can also improve in line with improvements in country averages without any reduction in the level of their relative disadvantage. In other words, the circumstances under which the gap is reduced and the trajectories each nation takes in addressing wide academic disparities at the bottom do matter as this directly influences the interpretation of results.

4.1.4 Tracking academic progress and bottom-end achievement gaps through PISA rounds

The relationship between the bottom-end achievement gap and academic progress between PISA cycles is certainly non-linear. Moreover, in the majority of countries, the trajectory varies according to the subject. The intricate dynamics of this relationship (sometimes presented as U-shape) is presented below on some national examples grouped under a few categories.

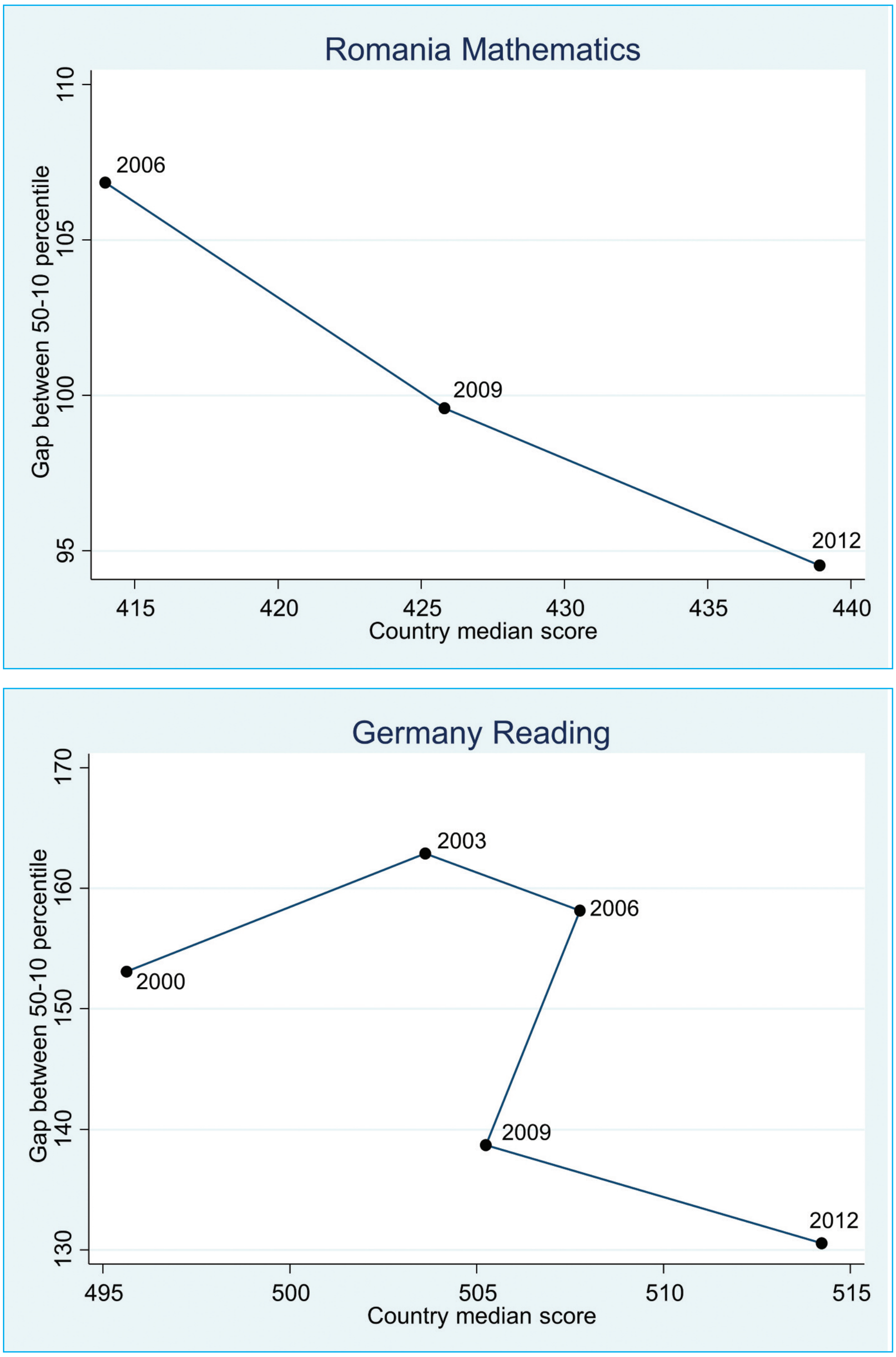
Academic progress with the focus on low-achievers. Some countries have managed to improve average test scores as well as reduce the achievement gap at the 'bottom' over the years. Germany and Romania are two countries with very different paths for improvement but similarly impressive results (Figure 7 page 28). Academic progress is imperative for a country like Romania, which has median scores substantially lower than the OECD average in each of the three subjects. Over the years the country managed to improve median scores alongside a further reduction in the already small (compared to the OECD average) relative disadvantage. Romania's median scores in mathematics were up by 25 points between 2006 and 2012 while at the 10th percentile the improvement was 37 points. Mexico and Turkey also showed similar targeted progress.

Germany's path for improvement was not as linear as in Romania, but it lifted up achievement in reading and reduced relative disadvantage through strong support of its academically weakest students. From 2000 to 2012, the country's median scores in reading increased by 19 points. During the same period, the group of low achievers at the 10th percentile improved by 43 points. In mathematics, the group at the 10th percentile made a three-fold improvement compared to the median (a change of 22 points and 7 points respectively) between 2003 and 2012. A similar trend is found in science. Further trend points are needed (post 2012) to see if the changes will be sustained.

With varying magnitude, similar progress in reading was made in Denmark and Japan (between 2003 and 2012), Belgium (from 2006 to 2012), and between two recent rounds in Estonia, Ireland and Poland. Arguably, the positive trend for some of these countries is too short to make assertions about its sustainability in the long run. More in-depth analysis of national policy changes is needed to link the positive trend with the national educational process. Yet the results are encouraging, indicating that these countries can be on the right 'track' for an inclusive educational progress when low-achievers are 'taken along' and supported in their educational needs.

Increased relative disadvantage and slipped performance. Quite a contrasting picture is found in some countries where deteriorating academic progress was matched with widening bottom-end inequality. This can be illustrated through examples of the Netherlands in mathematics (2006-2012), and Finland in reading (2006-2012) (Figure 8). The Netherlands experienced a continuous slide in the median mathematics scores between 2003 and 2012. A similarly large increase in relative disadvantage is observed between 2006 and 2012. In Finland, this trajectory can only be observed in reading with some U-turn change in the trend after 2006. As we saw earlier (Figure 5), Finland experienced an increase of 19 score points in the bottom-end gap between 2006 and 2012. This was accompanied by a downwards slide in the median reading scores also by 19 score points. This implies a disproportionate fall in the scores of children at the 10th percentile. Indeed, between 2006 and 2012, the reading scores of children at the 10th percentile in Finland fell twice as much as at the median (41 points compared to 19 respectively). This is a significant fall given the likely 'floor effect' in test scores of lowest achievers. Since their scores are already quite low, they do not have much room to fall further even though PISA does not have a minimum score (a student cannot 'fail' PISA).

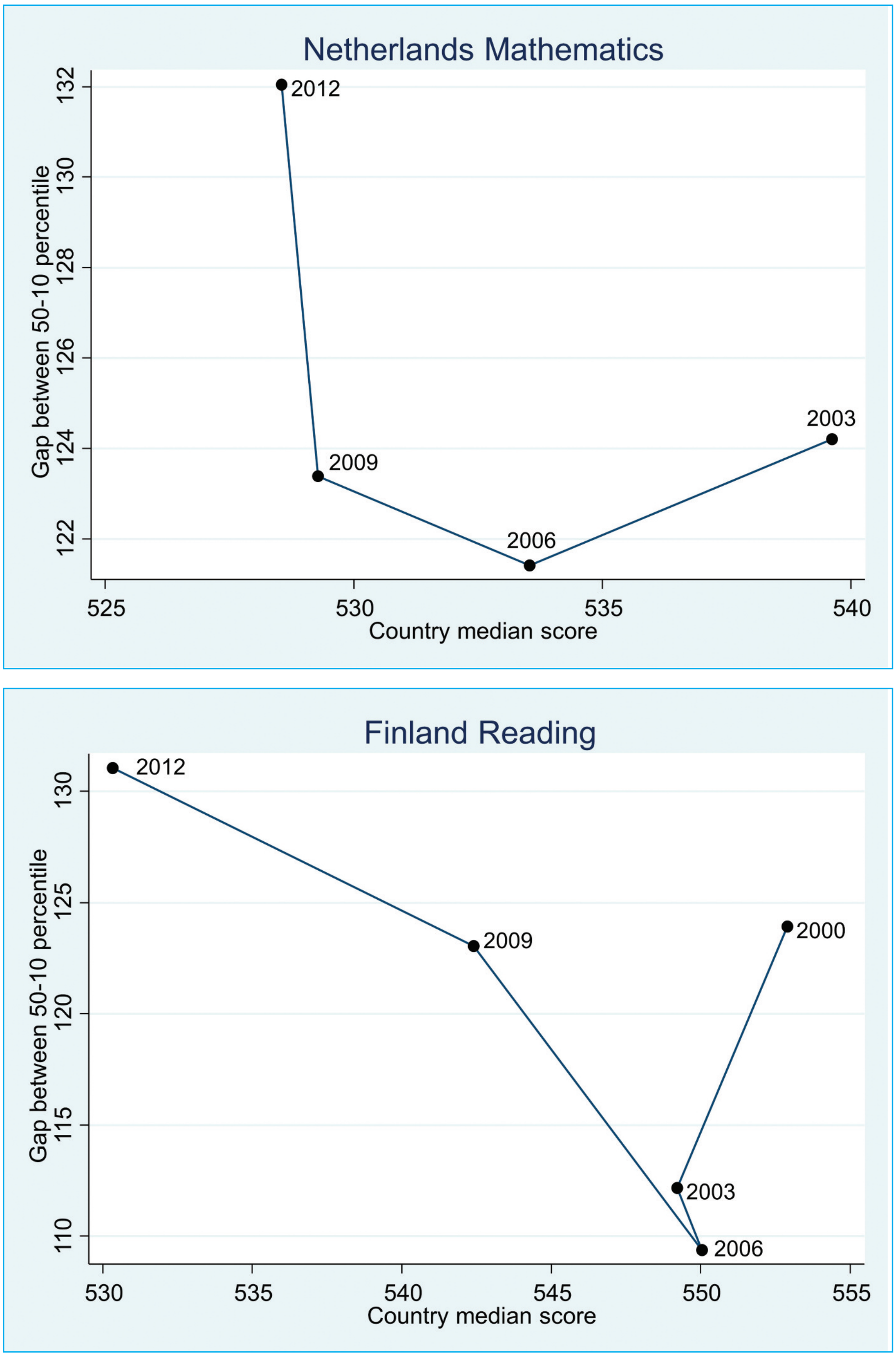
Figure 7 – ***Academic progress: increase in median performance and reduction in bottom-end inequality***



Note: Romania (mathematics): 2009 vs.2006, coef.-7.2, p<0.001; 2012 vs. 2006, coef.-12.3, p<0.000; Germany (reading): 2003 vs. 2000, coef. 9.8, p<0.001; 2006 vs. 2000, coef. 5.0, p<0.001; 2009 vs. 2003, coef. -14.4, p<0.001; 2012 vs. 2000, coef. -22.6, p<0.001
Source: PISA 2000/2003/2006/2009/2012

As well as the Netherlands and Finland, Sweden experienced a similar regressive trajectory in reading scores (2003-2006 and 2009-2012) and science (2006-2012), Canada in reading (2009-2012) and the Slovak Republic in mathematics (2009-2012).

Figure 8 – ***Deteriorating median performance and widening achievement gap***



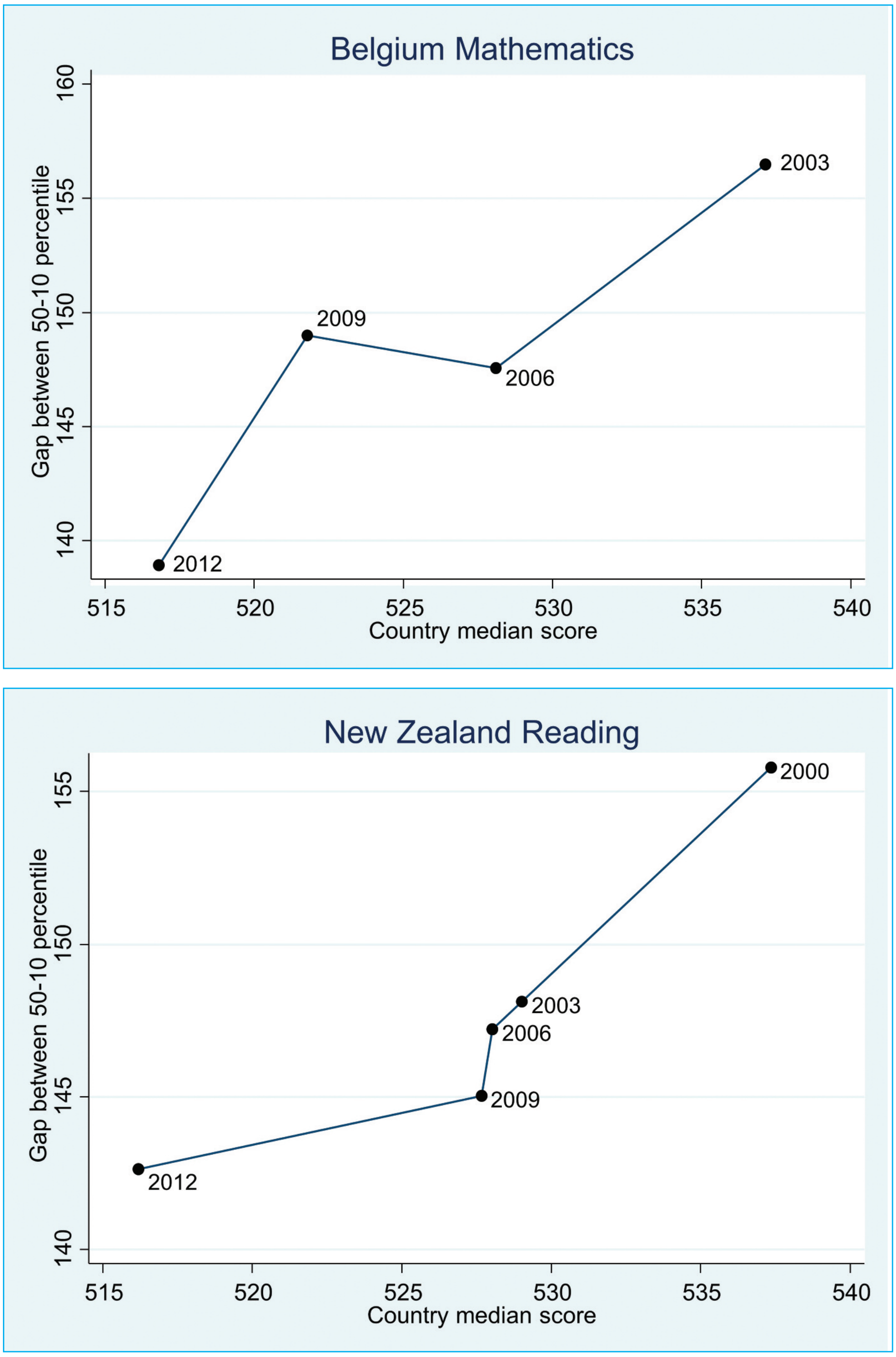
Note: The Netherlands (mathematics): 2006 vs. 2003, coef. -2.7, p<0.001; 2009 vs. 2003, coef. -.82, p<0.001; 2012 vs. 2003, coef. 7.8, p<0.001; Finland (reading): 2003 vs. 2000, coef. -11.8, p<0.001; 2006 vs. 2000, coef.-14.5, p<0.001; 2006 vs. 2003, coef. -.86, p<0.001; 2009 vs. 2000, coef. 7.2, p<0.001
Source: PISA 2000/2003/2006/2009/2012

'Double contraction' - reduced relative disadvantage but deteriorating average performance. These cases are of particular interest as they demonstrate that the reduction in relative disadvantage cannot be considered in isolation from the trend in overall academic progress. Here, the reduction of relative inequality is hardly a reason for celebration. Although the gap is narrowed, the achievement of the weakest students is likely to remain the same or get worse. We find that this type of trajectory is very common for mathematical achievement and is most often observed between 2003 and 2006 and between the last two rounds (2009 and 2012). Australia, Belgium, Denmark, New Zealand, Sweden and the United States, all have a similar zigzag trajectory across years in national mathematics performance. In Sweden, the reduction in the median maths score from 2003 to 2012 was 31 score points. It was supported by a similar drop in the scores at the 10th percentile (-27). Between 2006 and 2012 the fall in scores at the 10th percentile exceeded the median (by 27 and by 24 respectively). Belgium had experienced a different dynamic from 2003 to 2012 (Figure 9). Its scores at the 10th percentile declined but to a lesser degree than at the median (by 3 and by 20 respectively).

Finland, France, Greece, Hungary and Switzerland also experienced this double contraction at the lower end of the distribution and median score but only between the last two rounds. Very few countries have experienced this kind of development in reading with the exception of Australia (from 2000 to 2006 and from 2009 to 2012) and New Zealand (Figure 9 page 31), USA and Chile (from 2009 to 2012). In New Zealand the reduction of achievement at the 10th percentile was lower than at the median in all PISA survey rounds (by 8 score points compared to 21 from 2000 to 2012).

Low achievers are left behind. Finally, there are country cases when academic progress is not accompanied by narrowing relative disadvantage. As the educational system shifts average scores higher up, low achievers can be said to be left behind. We find this to be a rare scenario of academic development for the sample of countries we studied. Korea is the nation where this path is most noticeable (Figure 10 page 32). In 2012 Korea was at the top of our league table in mathematics with a median score of 557. Between 2003 and 2012 country's average score in mathematics improved by 12.4 score points but among children at the 10th percentile it constituted only 2 score points. Between 2009 and 2012, improvement in the median test score (by 8 score points) coincided with a fall of scores of children at the 10th percentile (by 5 score points). In Portugal, children at the 90th percentile improved their maths scores by 30 score points between PISA 2003 and 2012. Those at the 10th percentile also improved but by only 11 points. So, clearly, academic excellence was not equally distributed among all children in the country.

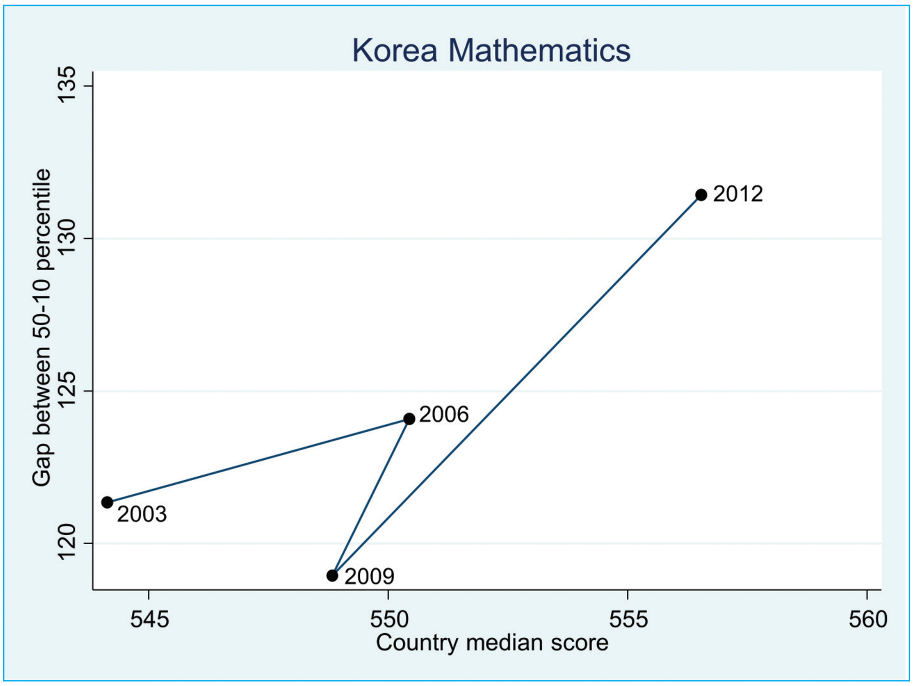
Figure 9 – ***Decrease in median score and bottom-end achievement gap***

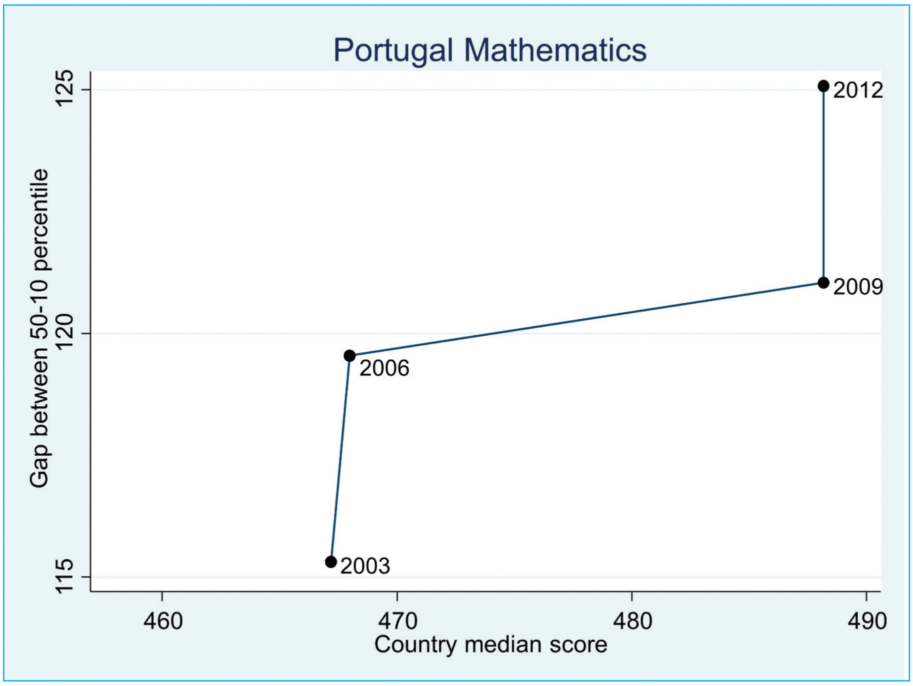


Note: Belgium (mathematics): 2006 vs. 2003, coef. -8.9, p<0.001; 2009 vs. 2003, coef. -7.5, p<0.001; 2012 vs. 2003, coef. -17.5, p<0.001; New Zealand (reading): 2003 vs. 2006, coef. -7.6, p.<0.001; 2006 vs. 2000, coef. -8.5, p<0.001; 2009 vs. 2000, coef. -10.7, p<0.001; 2012 vs. 2000, coef. -13.2, p<0.001.
Source: PISA 2000/2003/2006/2009/2012

Overall, our results do not support a notion of inevitable 'trade off' between higher academic achievement and inequality in educational outcomes. We also see that the trajectories of academic progress and relative disadvantage do not always coincide. Specifically, narrowing the gap can happen without academic improvement among low achievers but rather at the expense of deteriorating median academic performance.

Figure 10 – ***Academic progress in mathematics with widening bottom-end acheivement gap in Korea***





Note: Korea (mathematics): 2006 vs.2003, coef. 2.7, p<0.001; 2009 vs. 2012, coef. -2.4, p<0.001; 2012 vs. 2003, coef. 10.1, p<0.001; *Source:* PISA 2000/2003/2006/2009/2012

4.2 Absolute educational disadvantage

A first look at the incidence of low performance in reading (Figure 11) reveals striking variations across the countries studied. But no particular geographic pattern emerged. The group of countries with the lowest proportion of children below proficiency level 2 (less than 12%) includes two countries from East Asia (Japan and Korea), one from North America (Canada), as well as four European nations (Estonia, Finland, Ireland and Poland). These are all countries with high median scores in reading. Countries with the highest levels of absolute disadvantage (above 25%) are dominated by educational systems from Latin America (Chile, Mexico) and countries of the Central European region, which went through major structural reforms in the early 1900s (Bulgaria, Romania, the Slovak Republic).

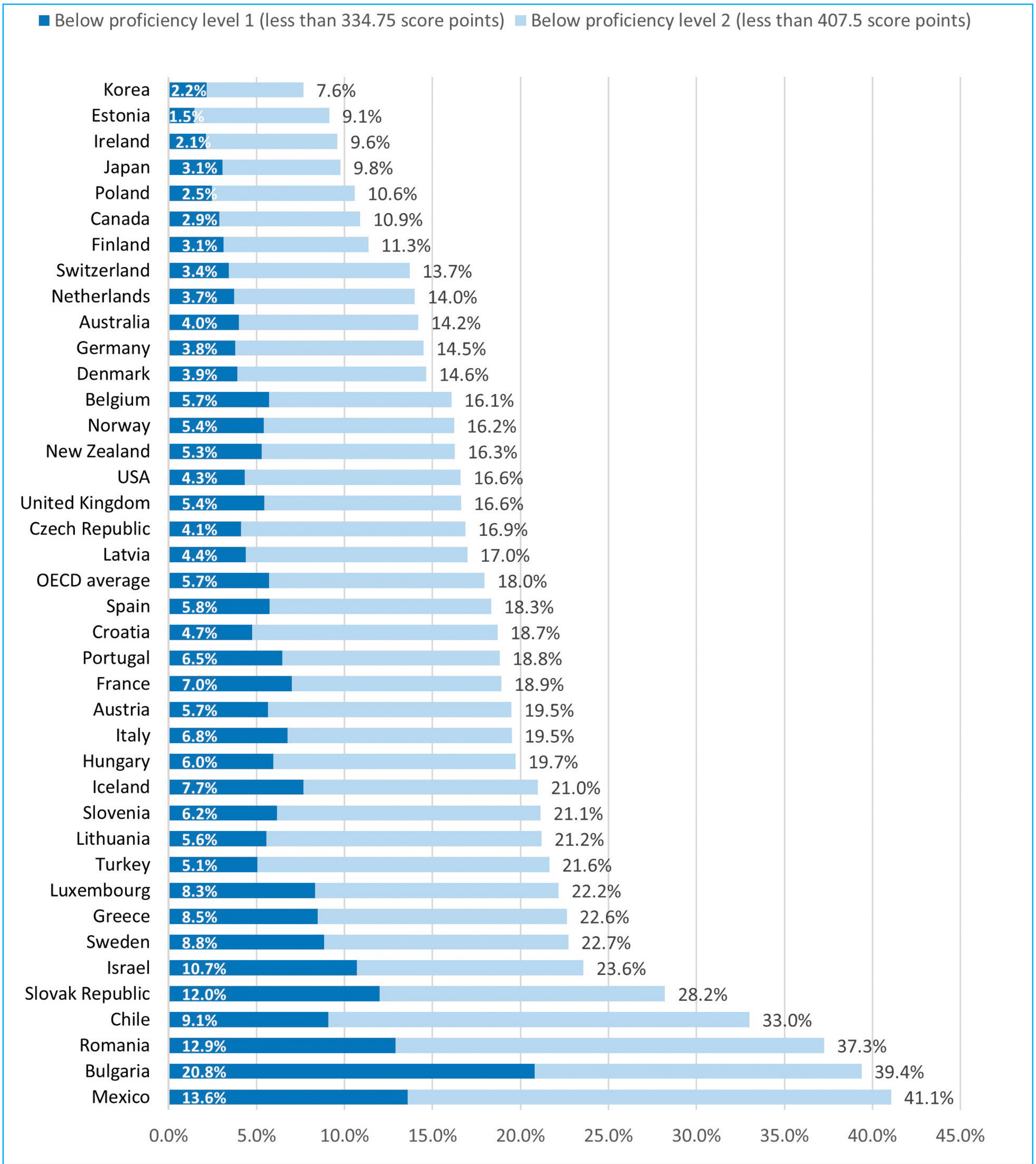
We find a clear and very strong negative correlation between median scores and the incidence of low performance in all three subjects (mathematics: r=0.97, p<0.001; reading: r=0.95, p<0.001, science r=.96, p<0.001). Yet, we see that with 23% and 22% of children below proficiency level 2 in reading, Sweden and Luxembourg rank lower than Turkey, despite having median reading scores 15 and 17 score points above those of Turkey. Although some sample bias cannot be ruled out, we do not have grounds to claim a 'trickle down' effect from high average achievement to reduce low performance indicating that a specific set of policies might matter too.

Children falling below proficiency level 2 in reading are not a homogeneous group in terms of acquired basic literacy skills. The group includes 15-year-olds at the proficiency level 1A (defined by PISA as scores just below level 2 but higher than level 1B[12]) and students below level 1B (less than 334.75 score points). Hereafter we distinguish these two groups as being low performance[13] and severe low performance, implying that the latter group has greater learning needs. For example, a student at the lowest level 1B in reading has to locate a single piece of explicitly stated information in a simple text with a familiar context. Required interpretation relates only to simple connections between adjacent pieces of information in the text. A student working at the higher level 1A in reading would have to recognise more than one independent piece of information, the main theme and author's purpose and make interpretive connections between information in the text and common everyday knowledge. The correlation between the proportion of children below level 1 and the country's median scores in reading and science is somewhat weaker than previously observed for 'below proficiency level 2' (reading r=80, p<0.001, science r=0.83, p<0.001, mathematics: r=.93, p<0.001,).

Figure 11 (page 34) presents the composition of low performance (below proficiency level 1 and below proficiency level 2) across 39 countries. The proportion of children below level 1 within the group of all low performers varies between countries so that the ranking would change if we were to take 'below proficiency level 1' as our guide. Ignoring these differences within the low performance group would risk some underestimation of the 'depth' or severity of educational need within a national context.

[12] Only reading makes a distinction between level 1A and 1B.
[13] This is consistent with the PISA definition (PISA, 2012).

Figure 11 – ***Proportion of children below proficiency level 2 and 1 in reading in the total population of 15 year olds in 2012***



Source: PISA 2012.

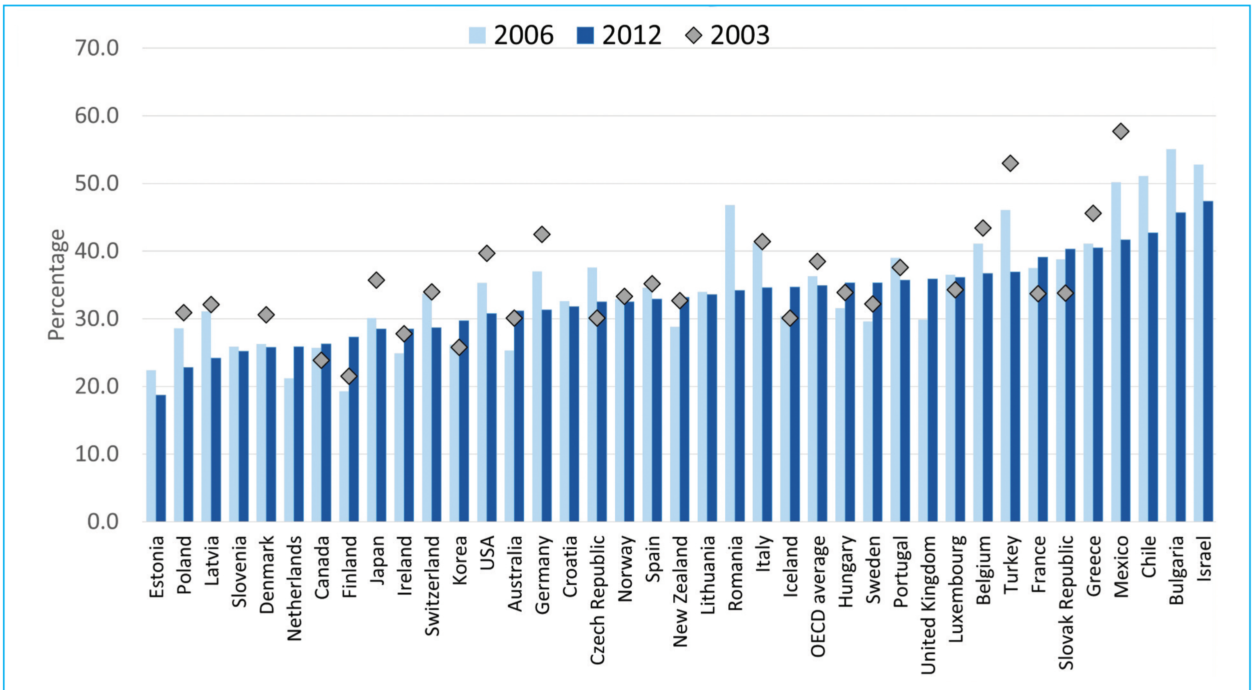
4.2.1 The 'depth' or severity of absolute educational disadvantage

The depth of educational disadvantage can be seen as analogous to the poverty gap used for the analysis of poverty. Just as the incidence of poverty tells us little about the extent of income/consumption shortage in the household, the incidence of low achievement (the proportion of children below proficiency level 2) is not very helpful in identifying the severity of skills deficiency experienced within this group.

A simple way of presenting the severity of absolute educational disadvantage is through the proportion of 15-year-olds below proficiency level 1 ('lowest performers') in the group of children below proficiency level 2 ('low performers'). The graph in Figure 12 presents the absolute levels of this measure in mathematics across three PISA rounds. On average across OECD countries, lowest performers account for 32% among low performers. Unsurprisingly, countries that consistently score worst on both – proficiency level 1 and 2 – have the highest concentration of lowest performance, relative to the group of low performers, with Bulgaria and Israel showing the highest concentration of severe low performance.

Between 2003 and 2012 or 2006 and 2012[14] the majority of countries reduced the severity of absolute disadvantage. The ratio declined among the top performing countries (Estonia, Poland and Latvia) as well as the bottom ones (Bulgaria, Chile, Israel and Mexico). At the same time, eight educational systems increased the concentration of severe educational needs (Finland, France, Iceland, Korea, the Netherlands, the Slovak Republic, Sweden and the United Kingdom) between 2006 and 2012.

Figure 12 – ***Proportion of children below level 1 in relation to the group of low performers (below proficiency level 2) in mathematics***

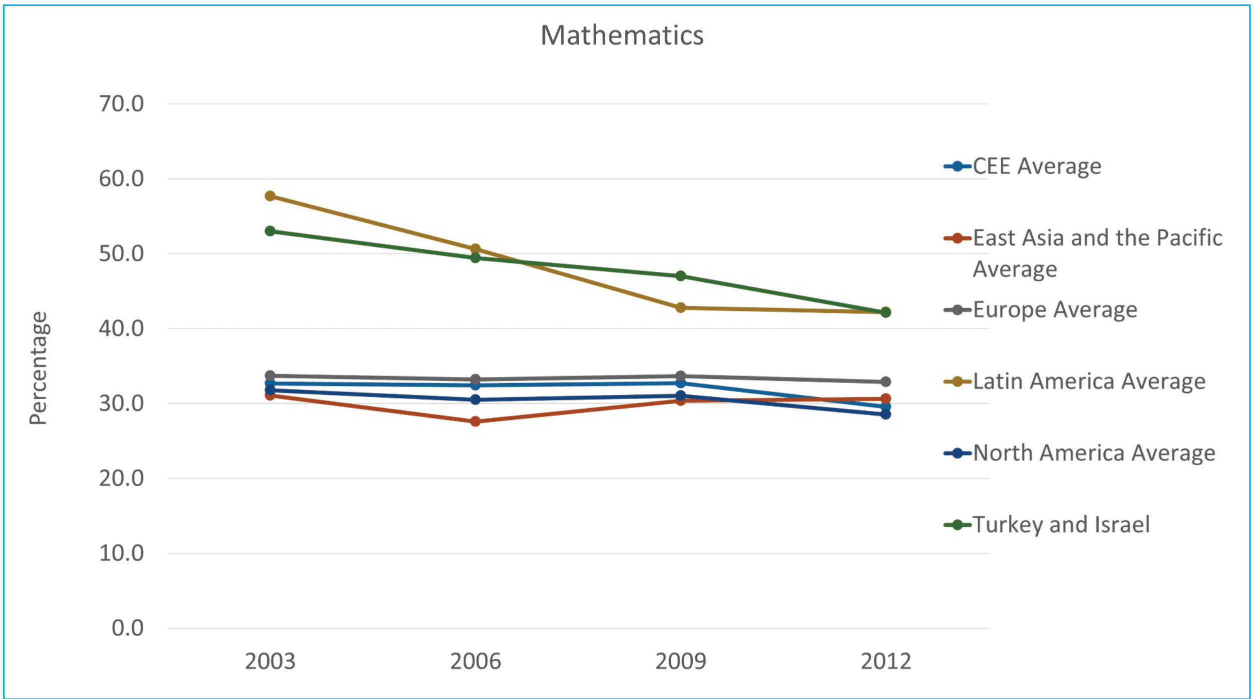


Note: Austria is excluded from trend analysis.
Source: PISA 2003/2006/2012.

[14] For countries which joined PISA in 2006.

The following highly aggregated view of trends in the relative share of severe absolute disadvantage (Fig. 13) across regions and PISA rounds provides a snapshot of the direction and the magnitude of change. Central European and Latin American countries have made remarkable progress in mathematics over the years of their participation in PISA. Although they still retain a high concentration of lowest performing students, compared to Western European countries and other regions, they are on track to close this gap. Middle eastern countries (represented by Turkey and Israel) as well North American countries also reduced the share of extreme disadvantage in the lowest performance group. In North America the change is driven by the United States which saw a 10 percentage points reduction in the ratio from 2003 to 2012.

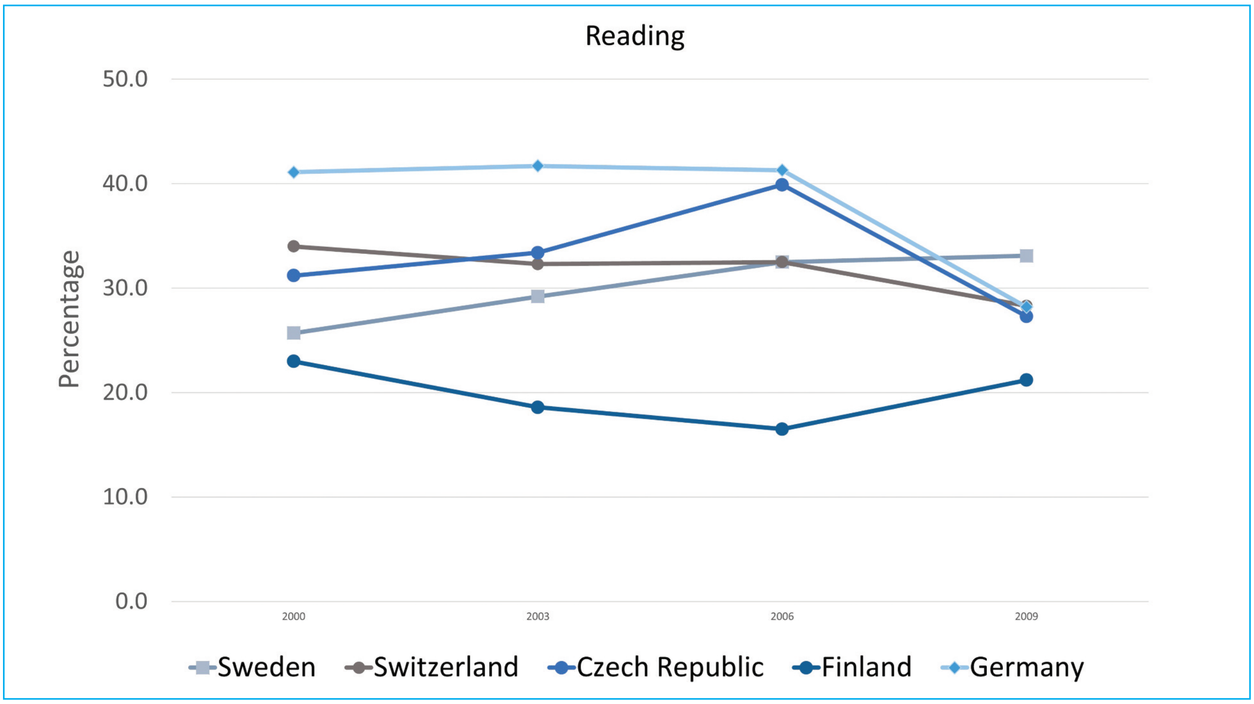
Figure 13 – ***Change in the proportion of severe low performance relative to all children falling below proficiency level 2***



Note: The CEE region includes Bulgaria, Croatia, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovak Republic and Slovenia. East Asia includes Australia, Japan, New Zealand and the Republic of Korea.
Source: PISA 2000/2003/2006/2009/2012

Given the complexity and variability of educational systems within each region it is not surprising that we find quite diverging trends among the selected countries. Figure 14 (page 37) shows the change in the severity of absolute disadvantage in reading in some Western European countries. Germany is one of the countries which achieved noticeable progress on this indicator over the years of its participation in PISA. In 2000, 41% of all low performers in Germany were children below level 1B in reading. By 2009 this was reduced to 13% and then further to 11% in 2012. Positive dynamics is also observed in the Czech Republic and Switzerland, in particular between 2006 and 2012. Somewhat diverging pathways are observed for Finland and Sweden. Both countries started in 2000 at a very similar level of severe absolute disadvantage: 26% in Sweden and 23% in Finland. In Finland, the positive trend was reversed between 2006 and 2012 when the concentration of lowest performance reached 28%, exceeding that of 2000. Sweden followed a consistent negative trend increasing the share of lowest performers (below level 1B) to 39% in 2012.

Figure 14 – ***Trends in severe low performance in reading in selected Western European countries***



Source: PISA 2000/2003/2006/2009/2012

This negative dynamic in the severity of absolute disadvantage in Sweden is troubling given it happened alongside a dramatic rise in the levels of low performing students in general (falling below proficiency level 2) in all three subjects: by 10 percentage points in reading (2000-2012), by 10 percentage points in mathematics (2003-2012) and by 6 percentage points in science (between 2006 and 2012). This indicates that the increase in severe low achievement happened with accelerated pace. This finding is generally in line with the OECD assessment that inequality in Swedish schools is on the increase (OECD, 2015). The drivers of deteriorating academic performance and rising educational inequality in Sweden have been actively discussed nationally as well as internationally (Blanchenay, Burns, & Köster, 2014; Sanandaji, 2014). The major educational reforms of the early 1990s – marketization in education, including the introduction of choice through voucher schemes – are among the most debated reasons linked to failures in the education sector. This is partially due to the evidence related to increased school segregations (West, 2014). At the same time, the complexity of the educational policy context means that other factors are at play as well, including the quality of teachers' training and methods of instruction in the classroom (Uppsala Universitet, 2014), or challenges of inclusive education of children with special education needs (Berhanu, 2011).

The levels of severe absolute educational disadvantage might be indicative of the priority placed on policies and provisions supporting extreme low performance. Although differences between individual students below proficiency levels 1 and 2 can be subtle,[15] it is reasonable to assume that as groups they might differ considerably in terms of risk of school drop-out, chances for further education or labour market success. Therefore, distinguishing them in terms of required learning support can be beneficial for the overall strategy to tackle absolute educational disadvantage.

4.2.2 The 'breadth' of absolute educational disadvantage as cross-subject low performance

So far, we have analysed the incidence and the depth of low achievement under proficiency levels 1 and 2 independently for each of three subjects: mathematics, reading and science. Yet, arguably, the groups of children who underperform in one subject are likely to have different characteristics academically and otherwise than those who systematically underachieve in all three subjects. Evidence indicates that systematic low-achievement in school is a complex and multifaceted phenomenon, with driving factors that are likely to vary across countries. Overall it might be an interplay of individual factors such as student motivation (Alderman, 2008), family background (Cassen & Kingdon, 2007) and, increasingly, school level characteristics (OECD, 2012).
Given the cumulative nature of education and learning, children with cross-subject educational difficulties are likely to be among low achievers for some time. Improving their performance to higher proficiency levels would require focused cross-curriculum support and possibly interventions of a systemic nature.

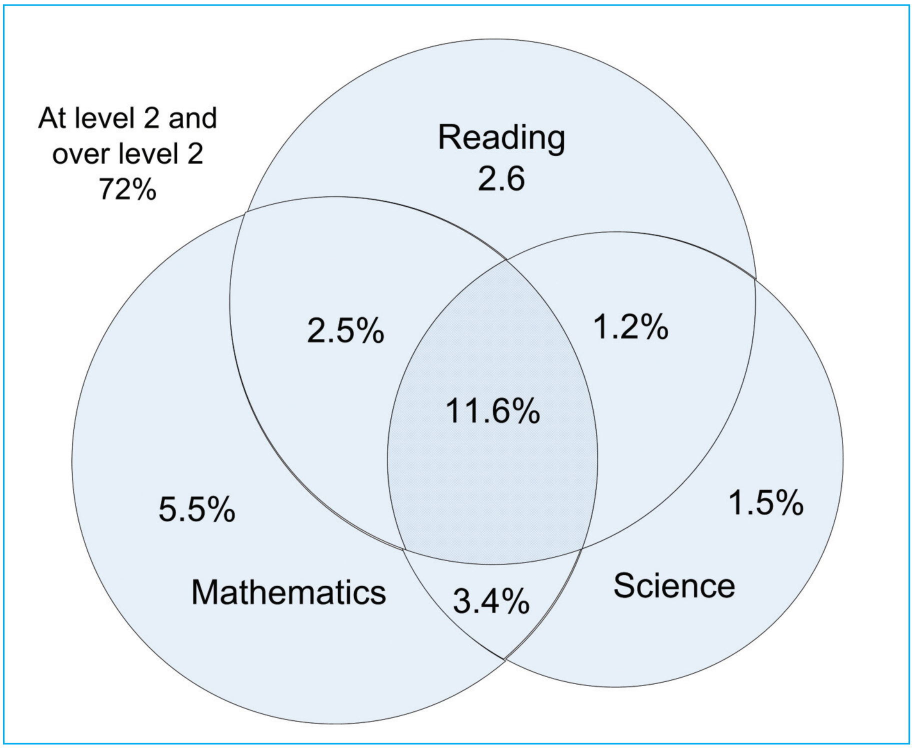
Figure 15 (page 39) presents overlapping low performance across mathematics, reading and science. On average, across OECD countries, 28% of 15-year-olds fall below proficiency level 2 in at least one of the three subjects. By far the largest group among them is children who score below proficiency level 2 threshold in all three subjects. Multiple or cross-subject low performance is indeed a prevailing trend in practically all countries studied. Estonia and Turkey are the only two countries where the overlapping sub-group is marginally smaller than the group with low performance only in mathematics.

Mathematics can be seen as one of the main drivers of low performance. The proportion of children below level 2 in maths (but not in other subjects) in 2012 was consistently higher than in reading and science. The OECD average of 5.5% is almost double that of reading (2.6%) and almost triple that of science (1.5%). The exceptions are Austria, Slovenia and Switzerland, where the proportion of 15-year-olds falling below level 2 in reading is higher than in maths (4.56%, 6.27%, and 3.15% respectively).

A sensitivity test confirms a somewhat equal subject association with the multiple overlap category in the majority of countries.[16] At the same time, low competency in maths as well as science is more common than overlap between any of these two subjects with reading. Among countries with the highest overlap of low performance in science and mathematics are Mexico (9.75%), followed by Turkey and Chile (8.2% and 7.7%). On the opposite scale are Estonia, Finland, Japan, Korea, and Slovenia with about 1% of children falling below proficiency level 2 in both subjects.

[15] We referred to the continuity in PISA proficiency level scale in our Data and Methods section.
[16] $r=0.57$ with mathematics $p<0.001$; $r=0.63$ with reading $p<0.001$; $r=56$ with reading $p<0.001$ for a pooled sample.

Figure 15 – ***Overlap of low performance (those below proficiency level 2) in mathematics, reading and science on average across OECD countries in 2012***



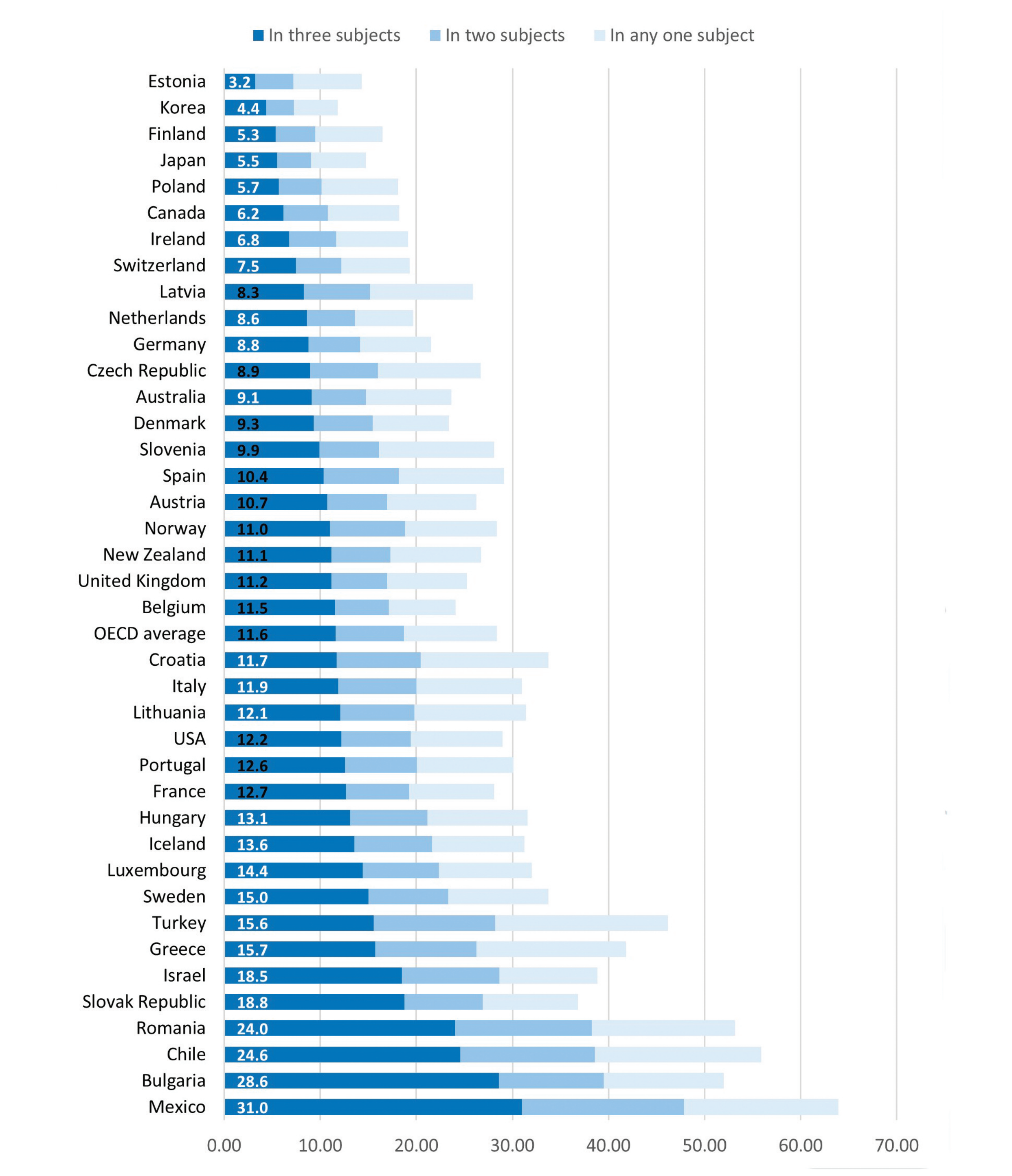
Source: PISA 2012

A more aggregated look at this measure is presented in Figure 16 (page 40) . Estonia tops the league table with only 3.5% of children falling below level 2 in all three subjects. On the other end of the scale is Mexico where the proportion of 15-year-old students scoring below proficiency level 2 in all three subjects is almost nine times higher (31%). This is one third of the enrolled students' population. Mexico is followed by Chile with 24% of children falling below level 2 in all three subjects and 14% in mathematics. Together with low achievement in the other subjects this leaves only 44% of children above level 2.

The evolution of cross-subject low performance over the years of countries' participation in PISA reflects not least their policy choices and agendas with respect to the support of low-performing students in multiple subjects. We hypothesised that educational systems can take three different routes in their progress in reducing cross-subject low performance. The first route is an improvement with the incidence of cross-subject low performance declining to a greater extent than in other sub-groups. The second route is an inclusive improvement with equal reduction in all low performing groups. The third route would leave the group of children with cross-subject difficulties behind (unequally shared rise/decline in the incidence of overlapping low performance). Below we present our findings under each of these categories.[17]

[17] Due to the comparability between science standardisation scales, the earliest PISA round which can be used for this analysis is 2006. The trend only has three observation points over a seven-year period. We consider this sufficient for our task of providing evidence on the most recent developments relevant for the current policy context.

Figure 16 – ***Overlap in absolute educational disadvantage in 2012***

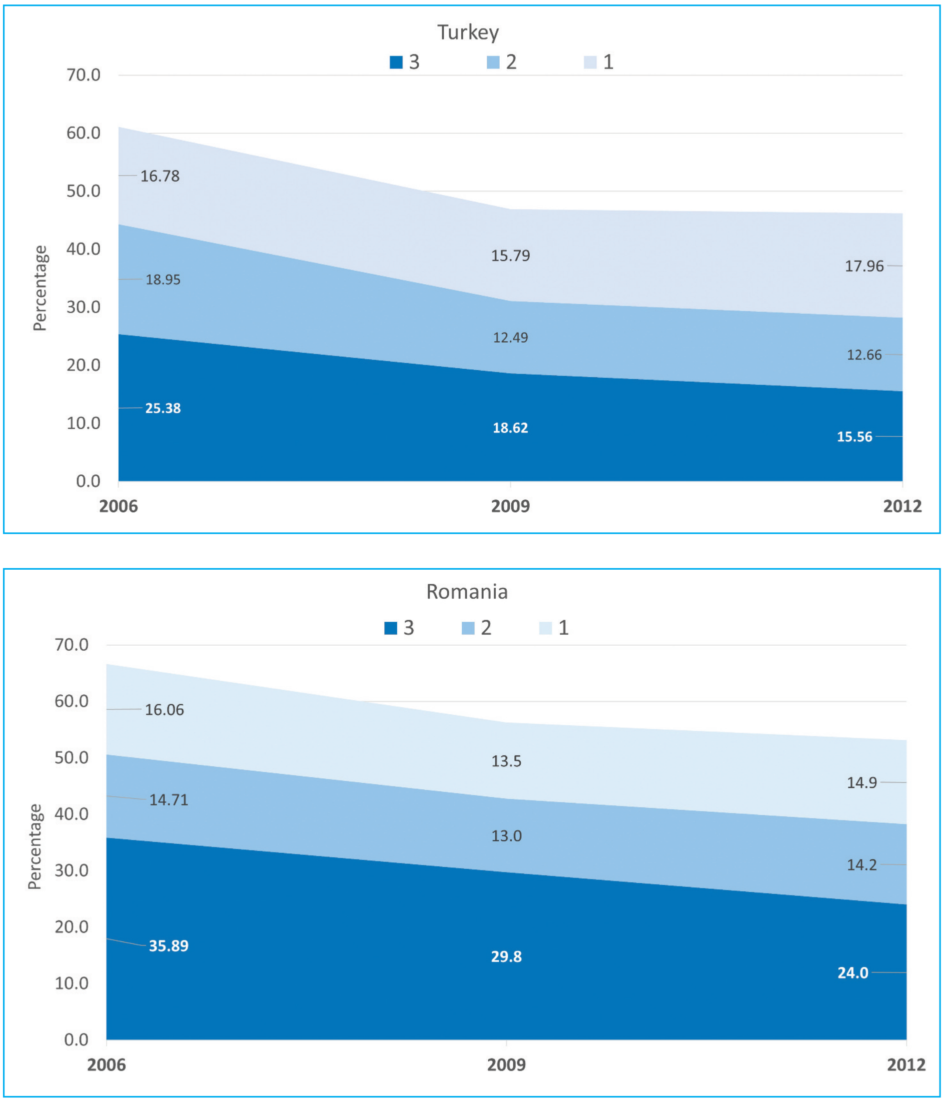


Note: Bar labels of countries in white indicate a statistically significant difference from the OECD average.
Source: PISA 2012

■ *Route A: Progress with a focus on the most educationally disadvantaged*

This path can be illustrated by the cases of Turkey and Romania (Figure 17). Having one of the highest levels of low performance among the countries studied, Turkey made the most noticeable progress in reducing cross-subject low performance. Between 2006 and 2012 it reduced the proportion of 15-year-olds with low performance in all three subjects by 40 percentage points, in any two subjects by around 33 percentage points. Meanwhile, subject specific low performance remained practically at the same level. These findings are consistent with PISA's assessment of Turkey's success in improving low performance among children with the greatest educational needs. According to a PISA 2012 report, low achievement in Turkey was reduced across all types of schools, with disparities in mathematical performance within schools narrowing between 2003 and 2012. Moreover, as no significant change is found in the scores of the top performing quantile, academic improvement is likely to be attributed to support targeted at low performance (PISA, 2014a).

Figure 17 – Targeted reduction of cross-subject performance



Source: PISA 2006/2009/2012.

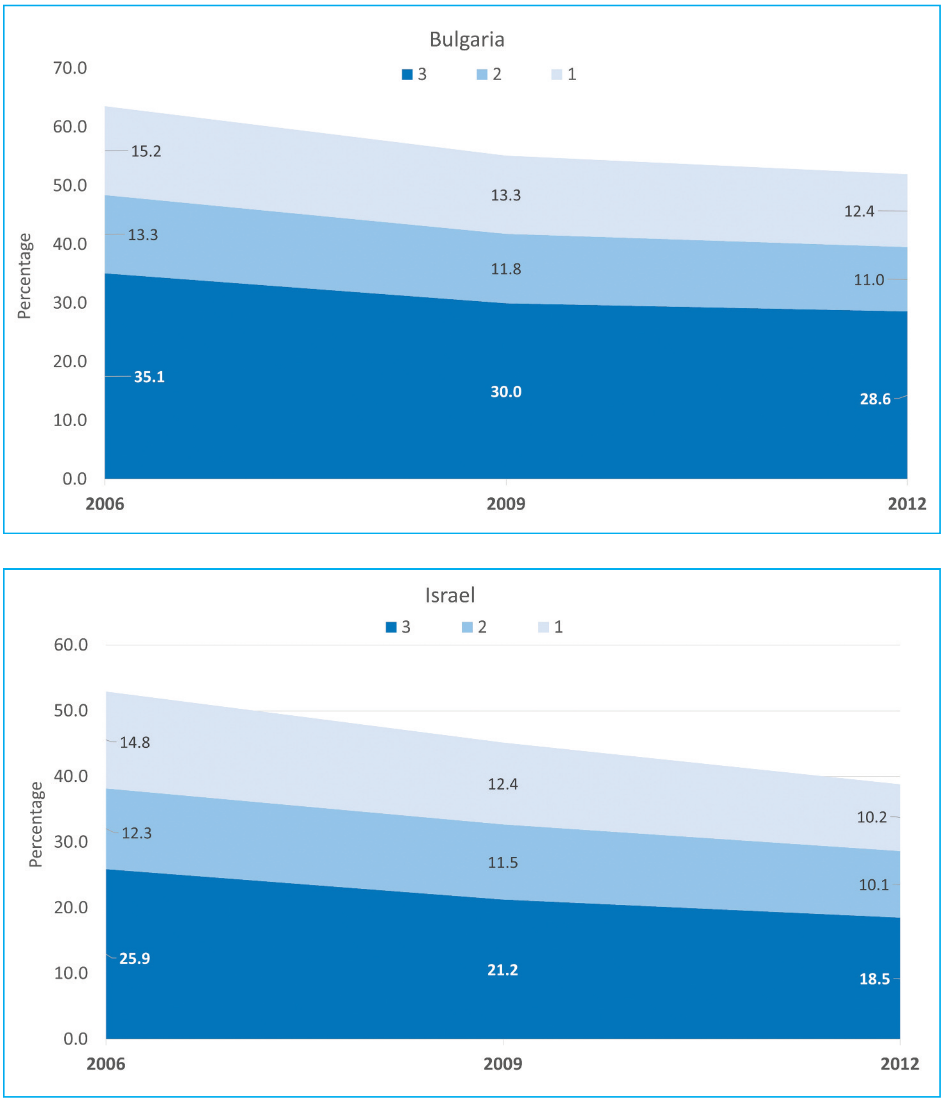
Romania has also made big leaps in catching up with other Central European countries over the years of its participation in PISA. From 2006 to 2012, the proportion of 15-year-olds below proficiency level 2 in Romania fell by 6 percentage points in mathematics, 16 percentage points in reading and by almost 10 percentage points in science. Our analysis indicates that this progress has been achieved thanks to the reduction in the group of low-performers in all three subjects. This group was reduced by 33 percentage points while other sub-groups remain practically constant. Estonia, Italy and Mexico are also worth mentioning here as they also saw a larger contraction in the group of children with multiple academic needs than in other sub-groups.

■ *Route B: Inclusive improvement*

We find that a considerable number of countries followed a more inclusive path of academic progress with the sub-group of low performers in all three subjects reduced at a similar rate as other sub-groups. For example (Figure 18), in Bulgaria the proportion of low performers was reduced

(continues on page 43)

Figure 18 – ***Reduction in low performance by sub-groups in Bulgaria and Israel***



Source: PISA 2006/2009/2012

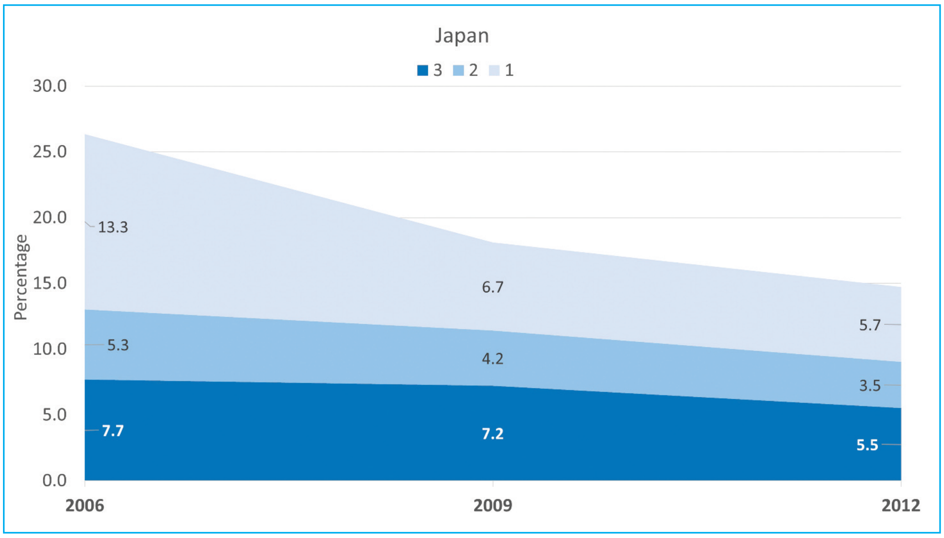
by 1.2 times (from around 15% to 12%) for students below proficiency level 1 in any one subject and by about 1.2 times among those who fail to reach level 2 in all subjects (from 35% to around 29%). In Israel, the reduction was about 1.4 times for these two groups and 1.2 times for children below level 2 in any two subjects. The magnitude and direction of change in these two countries reflect the fact that both Bulgaria and Israel have some of the highest levels of low achievement among the 39 industrialized nations studied.

Croatia, the Czech Republic, Italy, Latvia, Portugal and Switzerland are also countries where progress in the reduction of low performance affected all the sub-groups equally.

■ *Route C: Cross-subject low performers are left behind*

This is a route of a disproportionate change in the share of cross-subject low performance compared to other sub-groups. This might be the case when a) the academic progress (reduction in the incidence of low performance) is not shared equally with the most academically disadvantaged children or b) when the increase in the incidence of low performance affects the cross-subject group to a greater extent than other sub-groups. The case of Japan illustrates the first scenario (Figure 19) and Finland and Sweden the second (Figure 20 page 44).

Figure 19 – ***Reduction of low performance (falling below proficiency level 2) by subgroups of single or cross-subject academic difficulties***

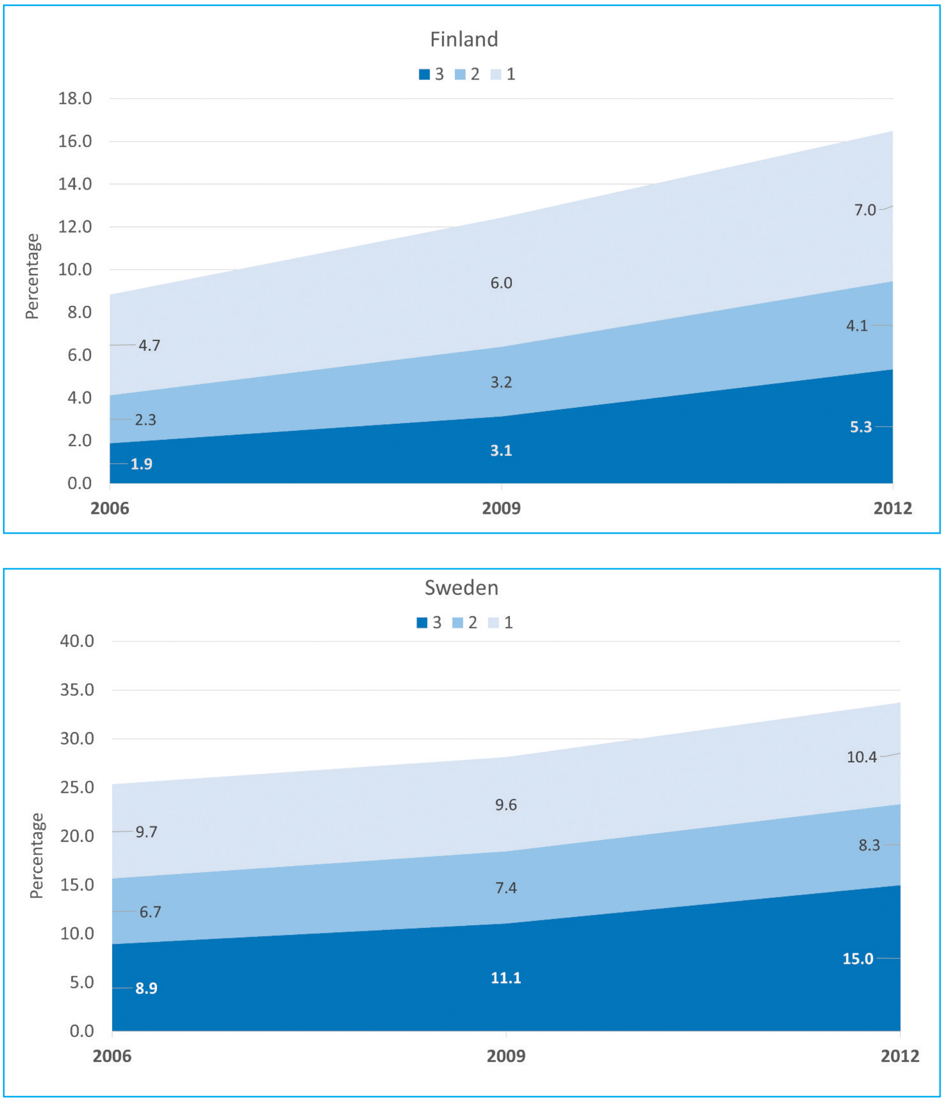


Source: PISA 2006/2009/2012

Japan showed consistent progress in the overall reduction of low performance in all three subjects from 2006 to 2012. The change affected all sub-groups but to a larger extent seems to be driven by subject-specific support. The group with multiple or cross-subject difficulties benefited from the progress to a much lesser degree, with a reduction of 28 percentage points compared to 42 percentage points in the group of children below level 2 in only one subject. As much as it is an unwelcome development for Japan, this is not a very common scenario across the 39 countries studied. Along with Japan, we find only Ireland has a similar pattern of leaving children with multiple educational needs somewhat behind while making strides towards academic progress for all.

A much more common scenario is the rise of cross-subject disadvantage when average academic performance shows signs of decline. Sweden and Finland are two countries where this deteriorating trend is most notable (Figure 20). Sweden saw a rather substantial negative change over the years with the share of 15-year-olds not achieving level 2 surging by 10 percentage points in reading (from 2000 to 2012), by 6 in science and by 9 percentage points in mathematics (from 2006 to 2012).
The rise in low performance disproportionately fell on the group with cross-subject low performance. The share of this group increased 1.7 times compared to 1.2 and 1.1 times for others respectively.

Figure 20 – ***Increase in low performance by sub-groups in Sweden and Finland***



Source: PISA 2006/2009/2012.

Finland also experienced a significant increase in low performance across all sub-groups.
In aggregate terms, the proportion of children below proficiency level 2 in mathematics increased by 6 percentage points, in reading by 5 percentage points during the same period and in science by 3.6 percentage points from 2006 to 2012. The group below proficiency level 2 in all three subjects increased to an even greater degree than in Sweden. Its incidence group increased almost three

fold (2.9 times) compared with a two-fold increase (1.8 times) in the sub-group of not proficient in any two subjects and a 1.5 increase for the sub-group of low achievers in any one subject.

The two countries pursued very different educational strategies over the last decade (Sahlgren 2015), so the similarity of this trend might not be indicative of the same factors that brought this change about. A wider set of socio-economic, ideological and other non-school factors might also act in intertwined ways with educational context. This does not, however, make it less important to monitor this negative trend in both countries.

5. COUNTRY RANKINGS: A COMBINED VIEW

Bringing together relative and absolute perspectives for a comprehensive view on inequality of educational outcomes is a challenging but policy relevant task. We draw on the analysis presented in the previous sections to present combined league tables based on two key indicators:

Relative educational disadvantage: score point difference between 50th and 10th percentile in mathematics, reading and science.

Absolute educational disadvantage: the percentage of children falling below proficiency level 2 in all three subjects.

To present the bottom-end inequality in education as an overall (system-wide) measure with an appropriate point of reference, each of these indicators were aggregated using z-scores in accordance with the methodological approach taken by the Report Card 9 (Currie et al., 2010). Achievement gaps in each subject were aggregated into inequality scores for each country. Aggregation is done by using standard scores (z-scores) which measure the standardized distance of any given country value from the group average (in this case OECD unweighted average). They are calculated by subtracting the country value (achievement gap) from the group average (in this case OECD) and then dividing the result by the standard deviation. The standardized scores in each subject are then averaged across three subjects to arrive at the overall score. Relative or absolute educational disadvantage 'close to average' is defined as a score within the range of -0.5 to +0.5[18] standard deviations from the OECD average. 'Relative or absolute educational disadvantage lower than the OECD average' is defined as having a standard deviation score greater than +0.5 from the OECD unweighted average. 'Relative or absolute educational disadvantage higher than the OECD average' is defined as having a standard score of less than -0.5 from the OECD unweighted average.

We use two different approaches in presenting the combined country rankings. The first approach replicates Report Card 4 (2002) in giving a two-dimensional view on inequality of outcomes (Figure 21 page 47). It compares average ranks in absolute and relative educational disadvantage. The countries whose educational systems have the lowest proportion of children scoring below proficiency level 2 in all three subjects (coloured in light blue) have the highest scores in absolute

[18] One SD in PISA is set at around 100 score points. For the sample size of around 30 countries, 0.5 SDs would be roughly equivalent to 1.96 SE (95% confidence interval).

disadvantage. Countries with the largest absolute disadvantage can be found at the bottom of the table (coloured in dark blue). Countries in the middle (-0.5 to 0.5) are those where differences are not substantially different from the OECD average.

There are marked differences between absolute and relative characteristics of inequality of outcomes across our sample of 39 nations. Chile and Romania are two countries with some of the highest levels of cross-subject absolute disadvantage but have a very small bottom-end achievement gap. At the same time, Germany and the Netherlands are ranked among the top performers in terms of absolute disadvantage, but marginally below -0.5 SD threshold under the relative measure.

We find a group of countries which are ranked consistently at the top of the league table on both measures. The educational systems of Estonia, Ireland, Latvia and Poland strike a balance between a small degree of bottom-end inequality and low incidence of cross-subject low-performance. At the opposite end consistently high relative, as well as absolute, disadvantage is found for the educational systems of Bulgaria, Israel, Slovak Republic and Sweden. These countries ranked substantially worse than the OECD average under both measures.

A matrix approach to presenting a combined view of relative and absolute disadvantage (Figure 22 and Figure 23 page 49) can potentially be more useful for stimulating debate on strategies for inclusive educational progress. Rather than focusing on countries at the 'top' and 'bottom', it differentiates educational systems between four groups:

- **Top right corner** – countries that tend to provide inclusive education with low levels of absolute and relative disadvantage. Estonia, Ireland, Latvia and Poland are substantively different from the OECD average on both measures. Estonia is noticeably ahead of other nations on the combination of these two indicators.
- **Bottom left corner** – countries which tend to have a high bottom-end achievement gap as well as a high proportion of cross-subject low performance. Bulgaria, Israel, the Slovak Republic and Sweden are substantially different from the OECD average on both measures. It is clear that Bulgaria's performance is an outlier even for countries in this group.
- **Top left corner** – countries which have a small achievement gap between the 50th and 10th percentiles, but high levels of absolute disadvantage. Chile and Romania are considerably distant compared to other OECD countries with z-scores around 2 below and above OECD average on both measures. This can be translated as around a 2 SD or 200 score point difference in achievement, or almost 5 years of schooling.
- **Bottom right corner** – countries with high relative disadvantage but a low proportion of cross-subject low performance. Germany and the Netherlands can be singled out here as those that are substantially different to the OECD average in z-scores on both measures.

Figure 21 – ***Combined league table on absolute and relative educational disadvantage***

	Absolute	Relative
Estonia	1.48	1.59
Korea	1.29	0.22
Finland	1.11	0.18
Japan	1.08	-0.48
Poland	1.05	0.79
Canada	0.97	0.28
Ireland	0.86	0.62
Switzerland	0.74	-0.12
Latvia	0.59	1.19
Netherlands	0.53	-0.70
Germany	0.51	-0.56
Czech Republic	0.48	0.30
Australia	0.45	-0.29
Denmark	0.40	0.66
Slovenia	0.30	0.46
Spain	0.22	0.36
Austria	0.16	-0.17
Norway	0.11	-0.28
New Zealand	0.08	-0.94
United Kingdom	0.07	-0.40
Belgium	0.01	-1.39
OECD average	0.00	0.00
Croatia	-0.02	0.88
Italy	-0.05	-0.26
Lithuania	-0.08	0.67
USA	-0.11	0.54
Portugal	-0.17	-0.10
France	-0.19	-1.36
Hungary	-0.26	0.15
OECD total	0.00	0.00
Iceland	-0.35	-0.46
Luxembourg	-0.50	-0.98
Sweden	-0.60	-0.61
Greece	-0.73	0.08
Israel	-1.22	-1.96
Slovak Republic	-1.27	-1.03
Romania	-2.20	1.77
Chile	-2.30	1.92
Bulgaria	-3.01	-0.97
Mexico	-3.44	2.19
Turkey	-0.70	1.76

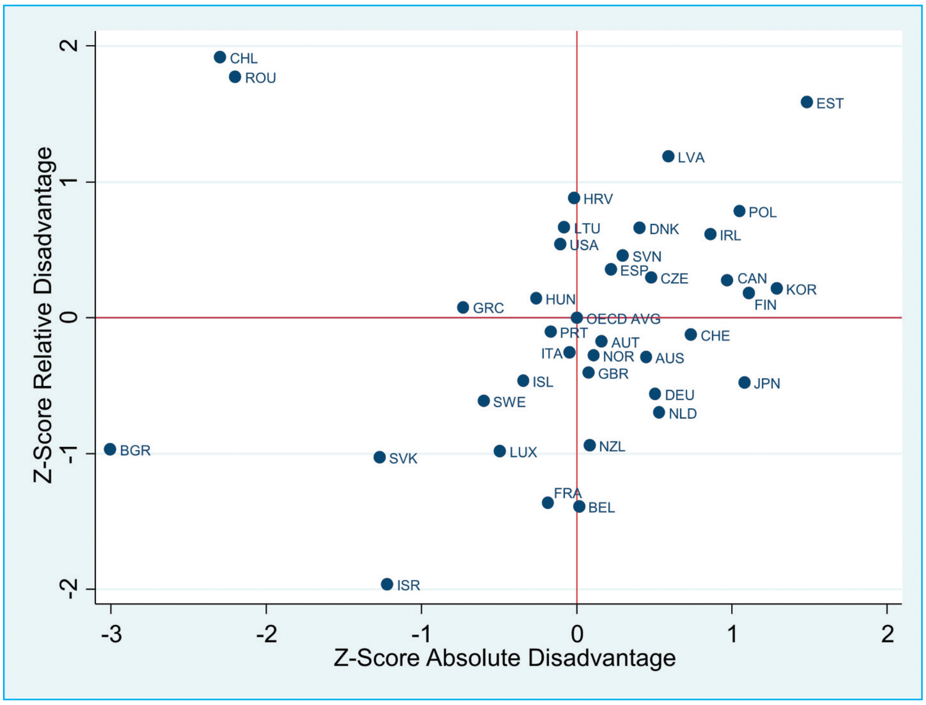
Source: PISA 2012

Figure 22 – ***League matrix on combined measures of relative and absolute educational disadvantage***

Low achievement gap **High** share of children below proficiency level 2 in all three subjects	**Low** achievement gap and **Low** share of children below proficiency level 2 in all three subjects
Chile, Romania, Greece, Hungary, USA, Lithuania, Croatia	**Estonia**, **Ireland**, **Latvia**, **Poland**, Korea, Finland, Denmark, Canada
Bulgaria, **Israel**, **the Slovak Republic**, Sweden France, Iceland, Portugal	Japan, Switzerland, Germany, the Netherland, Australia, Austria, New Zealand, Belgium, Great Britain, Norway
High achievement gap **High** share of children below proficiency level 2 in all three subjects	**High** achievement gap and **Low** share of children below proficiency level 2 in all three subjects

Note: Countries in bold are statistically different from the OECD unweighted average.
Source: PISA 2012

Figure 23 – ***Relative versus absolute disadvantage based on z-scores***



Source: PISA 2012

Considering the league matrix result in the context of our analysis of evolution of relative and absolute educational disadvantage provides a dynamic view on nations' progress for inclusive education for all. While some countries, notably Bulgaria, Chile and Romania are moving 'upward', catching up with a performance closer to the OECD average, others, such as Finland, France, Korea, Slovak Republic and Sweden demonstrate a worrying regressive trend in terms of providing support to their educationally disadvantaged children. The league matrix ranking should therefore be seen as an evolving 'landscape' which reflects the policy choices countries make about the needs of their educationally disadvantaged children and youth and the ways educational systems embrace diversity and equality of educational outcomes.

6 CROSS-COUNTRY VARIATION IN RELATIVE AND ABSOLUTE DISADVANTAGE: MACRO-ECONOMIC CONTEXT

International comparisons of PISA results need to acknowledge income differences across countries, since wealthier nations are able to spend more on education. The OECD records a positive relationship between mean maths performance in PISA 2012 and economic output per capita as well as spending per student (OECD 2014a: 35). Therefore, it is expected that median scores and the shares of students below proficiency level 2 in any subject are lower in richer countries as well as in those that spend more on education. It can also be posited that achievement gaps are lower in more equal societies because the strength of the association between PISA scores and parental socio-economic background tends to be higher in countries with greater inequality in disposable incomes (Causa & Chapuis, 2010). Evidence from macro-level studies also suggests that more and better quality education is associated with lower wage dispersion /inequality through the increased supply of skilled workers (OECD, 2011).

Table 1 – ***Correlations between PISA scores and macro-economic indicators (2012)***

	GDP per capita (PPP US$)		Gini coefficient		Spending on education, per student (PPP US$)	
	With Mexico and Turkey	Without Mexico and Turkey	With Mexico and Turkey	Without Mexico and Turkey	With Mexico and Turkey	Without Mexico and Turkey
Median score in maths	0.45**	0.37*	-0.61***	-0.50**	0.49**	0.36*
Median score in reading	0.44**	0.38*	-0.47**	-0.33	0.42**	0.31
Median score in science	0.38*	0.29	-0.51**	-0.36*	0.40*	0.25
50-10 gap in maths	0.44**	0.37*	-0.47**	-0.33*	0.43**	0.29
50-10 gap in reading	0.26	0.18	-0.27	-0.15	0.25	0.13
50-10 gap in science	0.58***	0.52***	-0.51***	-0.38*	0.56***	0.45**
% below proficiency level 2 in maths	-0.45**	-0.36*	0.64***	0.53***	-0.50**	-0.37*
% below proficiency level 2 in reading	-0.38*	-0.32	0.43**	0.29	-0.37*	-0.26
% below proficiency level 2 in science	-0.30	-0.20	0.47**	0.31	-0.32	-0.16
% below proficient level 2 in all three subjects	-0.34*	-0.27	0.49**	0.36*	-0.36*	-0.24
N	39	37	39	37	39	37

Base: OECD and/or EU countries in PISA 2012
Sources: GDP per capita in 2012 (International Monetary Fund, World Economic Outlook Database, April 2015); Gini coefficient in 2012 (Solt 2014 "The Standardized World Income Inequality Database (SWIID) Version 5.0, October 2014); cumulative spending on education in 2012 (OECD 2014, PISA Vol 1, Figure I 2.2).
*p<0.05; **p<0.01; ***p<0.001

Although disentangling causal relationships is far beyond the scope of this paper, we test these hypotheses on our key indicators of relative and absolute educational disadvantage using data for 2012. Table 1 reports bivariate correlation coefficients between the median scores, achievement gaps and shares of students below proficiency level 2 (in each subject and in all three subjects) and GDP per capita PPP US$, Gini coefficient and cumulative spending per student (thousand PPP US$).[19]

6.1 National income

Counterintuitively, countries with higher GDP per capita tend to have greater achievement gaps in maths and science (Figure 24 page 51). But the relationship is non-linear and there is more dispersion in achievement gaps among wealthier countries. Moreover, the effect of per capita GDP on the achievement gap in maths is no longer statistically significant when the median score is controlled for, suggesting that it may be spurious. The positive association is only observed because countries with higher median maths scores tend to have higher achievement gaps in the lower half of the distribution, while richer countries tend to have higher median maths scores. In contrast, the positive association of country income on the achievement gap in science persists even after controlling for the median score. Nevertheless, this is not evidence of a causal relationship, but merely an association observed among 39 countries.

6.2 Education spending per student

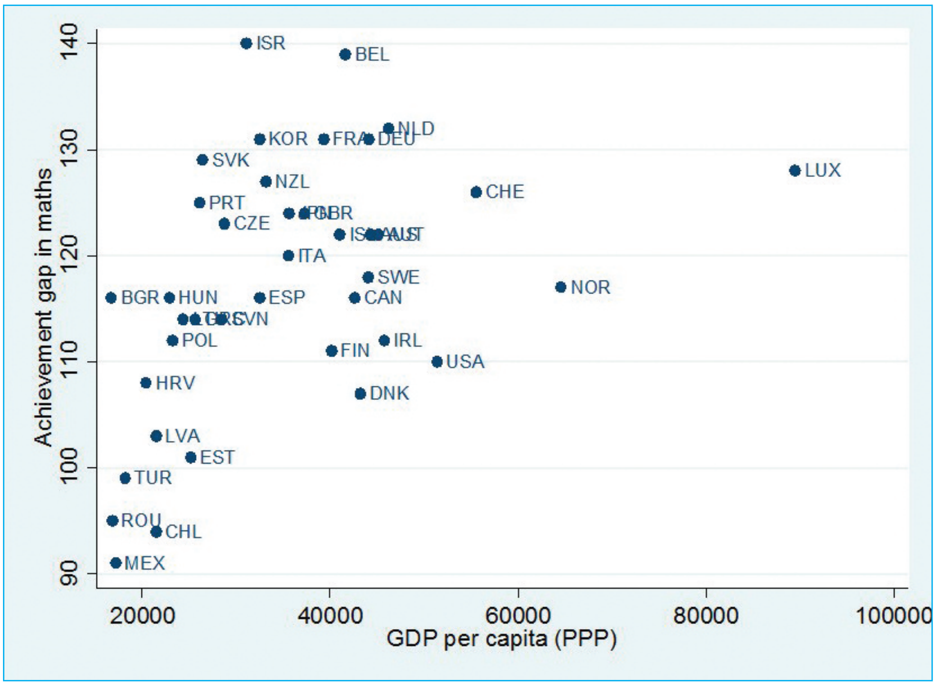
Median maths scores are significantly higher in countries with higher cumulative education spending per student (PPP US$), even if Mexico and Turkey are excluded (Figure 25a page 52). The relationship appears to be non-linear because Luxembourg registers middling median performance in maths in spite of the highest education spending by far per student in PPP US$. The effect of education spending remains large and statistically significant even after controlling for GDP per capita, which itself is no longer statistically significant.[20] This suggests that the level of spending per student, rather than country wealth per se, plays a crucial role in explaining the cross-country variation in median maths performance. In fact, education spending per student (and its quadratic term to allow for non-linearity) explains nearly one-half of the variation in median maths scores in PISA 2012 before controlling for school and student characteristics.

The achievement gap in maths appears to be larger in countries with more generous spending per student, but the correlation is no longer significant at $p<0.05$ if Mexico and Turkey are excluded (Figure 25b page 52). Moreover, the effect disappears if the median maths score is controlled for, indicating that the observed positive association may be spurious. However, robustness checks suggest that the positive association between the achievement gap in science and education spending is robust to both excluding Mexico and Turkey and controlling for the median science score. This does not suggest that higher spending inevitably leads to greater dispersion in science

(continues on page 53)

[19] Spending per student data is based on OECD estimations: expenditure is approximated by multiplying public and private expenditure on educational institutions per student in 2012 at each level of education up to the age of 15.
[20] There is a very high correlation ($r=0.95$, $p<0.001$) between GDP per capita (PPP US$) and spending per student (PPP US$) in 2012.

Figure 24 – ***Achievement gap in mathematics and science GDP per capita (PPP US$)***

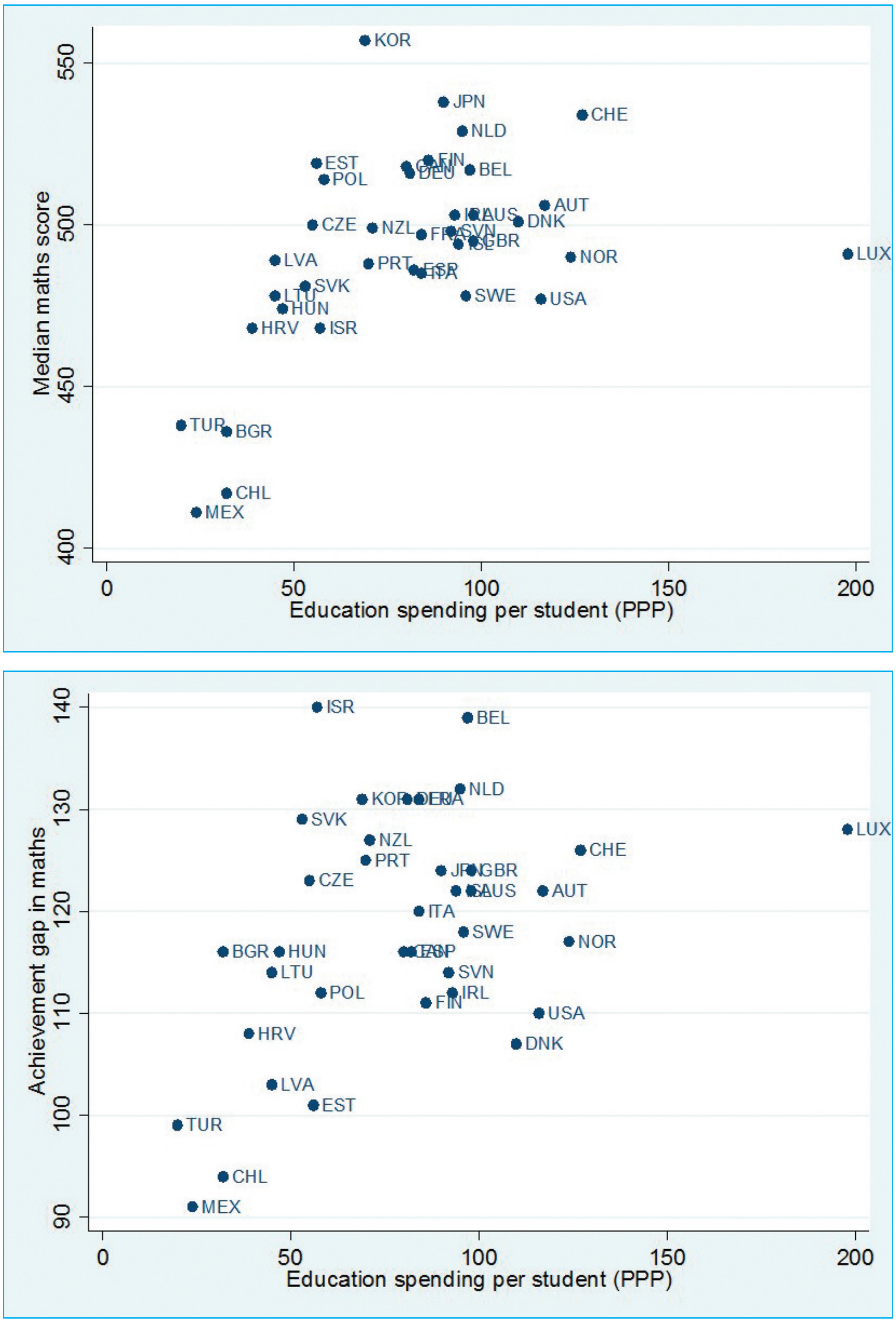


A. Mathematics scores

B. Reading scores

Sources: PISA 2012; GDP per capita in 2012 (International Monetary Fund, World Economic Outlook Database, April 2015).

Figure 25 – ***Median score and achievement gap in maths and education spending per student (PPP US$)***



A. Median

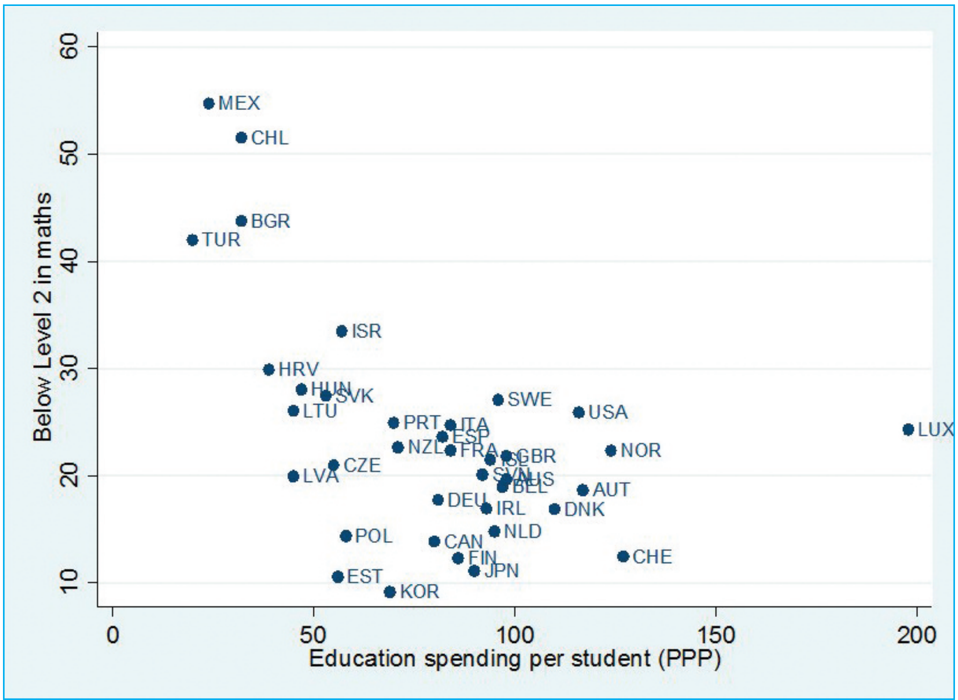
B. Achievement gap

Sources: PISA 2012; Spending on education in 2012 (OECD 2014, PISA Vol 1, Figure I 2.2).

scores in the lower half of the distribution, but merely that higher spending does not necessarily ensure greater equality in science performance. Indeed, countries that boast both higher median science scores and lower bottom-end inequality, such as Estonia and Korea, spend relatively little on education in absolute terms.

The proportion of children falling below Level 2 in maths tends to be lower in countries that spend more on education (Figure 26), although the effect is non-linear. It is robust to both excluding Turkey and Mexico and controlling for GDP per capita (which does not have a significant partial effect itself). Education spending per student (allowing for the non-linearity of its effect by including a quadratic term) explains half of the variation in the share of students performing below Level 2 in maths. At lower levels of spending this suggests that extra investment may lead to more students achieving the minimum proficiency levels, but returns are lower if spending is already relatively high. A similar pattern is observed for the cross-country association between education spending and the share of students falling below level 2 in all three subjects (not presented here), except that the association is somewhat weaker.

Figure 26 – ***Share of children below Level 2 in maths and education spending per student (PPP US$)***

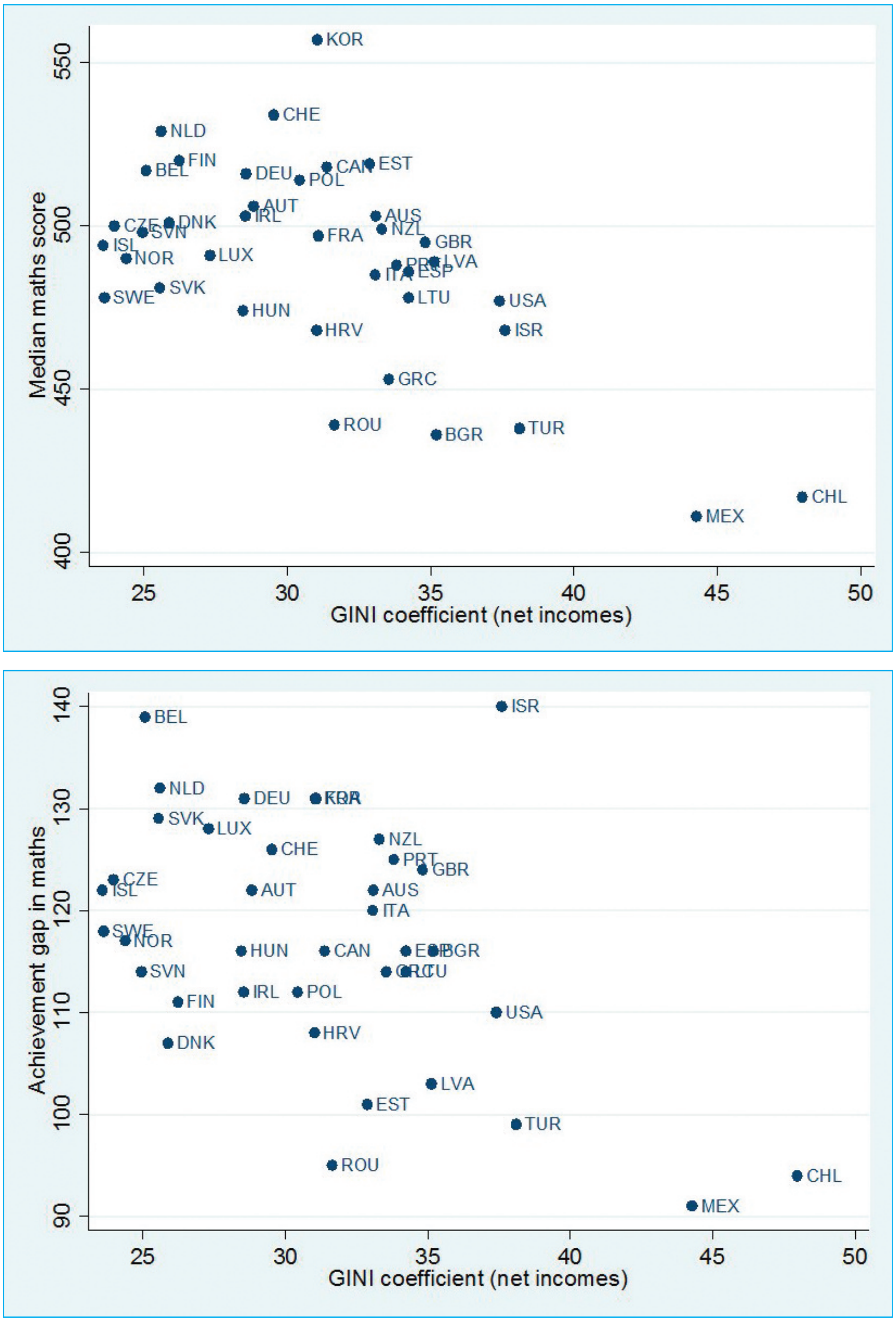


Source: PISA 2012, Spending on education in 2012 (OECD 2014, PISA Vol 1, Figure I 2.2).

6.3 Income inequality

Median maths performance tends to be lower in countries with higher inequality in disposable incomes as measured by the Gini coefficient (Figure 27a page 54). The effect is somewhat nonlinear, however: it is steeply positive at lower levels of income inequality (up to around 30%) and steeply negative at higher levels. Thus, countries with the most equal income distribution, such as Norway and Sweden, do not boast the best performance in maths, while the best performing countries, such as Korea and Switzerland, have low-to-average income inequality. The effect of the Gini coefficient is robust to controlling for total education spending per student. Income inequality and education spending (allowing for the non-linearity of their effects by including quadratic terms) explain two-thirds of the variation in the median maths scores in PISA 2012 before controlling for school and student characteristics (see Appendix).

Figure 27 – ***Median score and achievement gap in maths and the Gini coefficient***



A, Median

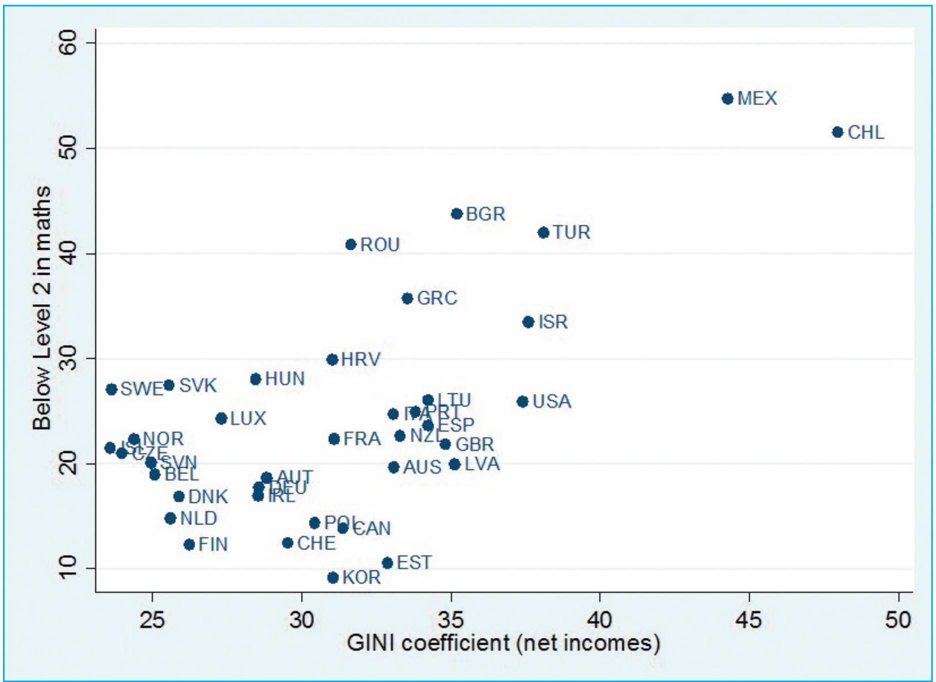
B. Achievement gap

Source: PISA 2012; Gini coefficient in 2012 (Solt 2014: The Standardized World Income Inequality Database (SWIID) Version 5.0, October 2014).

There is a suggestion of a negative association between the 50-10 achievement gap in maths and income inequality (Figure 27b), but it is driven by Chile and Mexico. It is also spurious: across all 39 countries, higher achievement gaps in maths are associated with higher median scores, while higher median scores are observed in countries with lower income inequality. A similar pattern is found with respect to science. There is no significant association between the achievement gap in reading and income inequality.

The share of students falling below proficiency Level 2 in maths is significantly lower in more equal countries (Figure 28). The positive association is robust to controlling for both education spending and per capita GDP. Yet, we find that the effect of income inequality is clearly driven by Chile and Mexico, the countries with highest income inequality and lowest share of students achieving Level 2 in PISA. The same pattern is observed for science and reading, as well as for the overlap between all three subjects.

Figure 28 – ***Share of students below Level 2 in maths and income inequality***



Source: PISA 2012; Gini coefficient in 2012 (Solt 2014: The Standardized World Income Inequality Database (SWIID) Version 5.0, October 2014).

Thus, achievement gaps in the lower half of the distribution tend to be higher in richer countries and those that spend more on education in absolute terms. However, this is largely an artefact of higher gaps being associated with higher medians and higher medians being observed in richer countries. Greater relative educational disadvantage is also observed in countries with more equal disposable incomes, but this too is mediated by median scores. It has to be stressed that most of these associations are non-linear and driven by a cluster of countries with low per capita GDP and low education spending, such as Bulgaria, Chile, Mexico, Romania, and Turkey. At the same time, these associations fail to explain how other countries with relatively low income and modest spending manage to maintain both excellent performance and low educational inequality, notably Estonia and Finland. High-quality comparable information on the structure of education systems would be necessary to fine tune this analysis.

Cross-country associations between macro-level indicators and absolute educational disadvantage, measured as the proportion of students falling below Level 2 in any subject and in all three subjects, all stack up in the expected direction. Fewer children fall below minimum performance benchmarks in richer and more equal countries, as well as those that spend more on education, but it is the latter that has the largest independent effect. Yet, it too would disappear if countries with the largest shares of low performing children were excluded.

7. DISCUSSION AND CONCLUSION

This study contributes evidence on the level and evolution of relative and absolute educational disadvantage across 39 industrialized nations. By focusing on bottom-end inequality of educational outcomes it hopes to stimulate debate on policy options related to the groups of children with greatest educational needs without losing sight of raising academic standards for all. Recognizing the aggregate nature of the analysis across a large and very diverse group of countries, it is not our intention to draw any prescriptive policy conclusions. As many studies pointed out, it is not advisable to base policy recommendations on PISA country rankings alone without the context of educational provision as a whole or understanding of the specific goals and the type of secondary schooling within a national context (Bulle, 2011). For example, continuity of educational provision means inequality in learning outcomes at the secondary level can be the direct result of problems at the primary school level (Choi & Jerrim. 2015). The latter was beyond the scope of this paper.

PISA collects data on a very specific sub-sample of the overall education system. So even though the analysis provides rich evidence on characteristics and the evolution of relative and absolute educational disadvantage, a complementary analysis of other surveys such as Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS) can provide a complementary test of the measures within different educational contexts or age group coverage. At the same time, one has to recognize the contextual as well as methodological differences between these educational surveys, particularly related to the item response models used, and resulting potential disagreements in country rankings (Micklewright & Schnepf 2006). Finally, as the goals of this paper were limited to measuring and describing the bottom-end inequality of educational outcomes of 15-year-olds, further examination of socio-demographic profiles of educationally disadvantaged children is required to determine factors that drive low achievement or low performance within national contexts.

This paper supported the premise previously found in the literature, that measuring inequality of educational outcomes in analogy with income inequality in relative terms is not appropriate due to problems associated with the standardisation procedures applied to educational data. We assert that the choice of indicators strongly depends on the corresponding definition of inequality and the research task at hand. We found that percentile range and standard deviation are both viable and consistent measures which produce very similar country rankings. But only the achievement gap as a range between 50th and 10th percentile corresponds to our research focus on 'bottom-end' inequality relative to the national median.

As with any country level indicator this measure is not without limitations. The interpretation of the absolute gap or its reduction at its face value should be taken with caution. Firstly, our analysis has shown that the bottom-end achievement gap cannot be considered in isolation from a) progress in median achievement and b) the scale and severity of academic disadvantage 'at the bottom' of the distribution. The view that a narrow achievement gap at the bottom indicates a positive outcome was challenged in our study by the cases of Chile, Mexico, Romania and Turkey. Having some of the smallest achievement gaps these educational systems have some of the lowest levels of median performance. The proposed complementary indicator which takes into account the scale of absolute educational disadvantage (absolute minimum of core skills and competencies in all three subjects) helps to place bottom-end inequality in the context of children's chances for human development compared to other nations.

The indicators of relative and absolute educational disadvantage at one point in time can be helpful for aggregate cross-country comparison. But without an analysis of the evolution of changes in key indicators it is easy to overlook the complex dynamics between national performance and bottom-end inequality, sustainability of progress or the magnitude of change. The paper stressed that any league table ranking has to be treated as an 'evolving landscape' where some countries show a worrying trend away from a relatively high overall position, while others might be well on track to catch up with other nations.

7.1 Relative educational disadvantage

We find marked differences in the scale of relative disadvantage between the countries in our sample. Translating them into years of schooling, based on PISA methodology, would imply a range from 2.5 years in Mexico, Chile, and Romania to 3.5 years of schooling in Belgium (mathematics) or Bulgaria (reading). But with even the smallest gaps amounting to a staggering 2.5 years of schooling, the question of 'top' and 'bottom' performing educational systems is hardly relevant. We argue that an average disparity of three academic years found across the OECD countries is unacceptably high. It imposes risks of marginalisation of children with the greatest educational needs, excluding them from many activities and social functions considered the norm for their age group in the society. Rather than using the OECD average as a benchmark, the countries can learn from the experiences of other educational systems like Estonia, where low relative educational disadvantage is achieved alongside very high median scores. Assuming that every educational system is unique, an identification of factors and areas which drive the change at the bottom of the academic distribution would need to be the first step in this direction.

Evidence on the dynamic relationship between the bottom-end achievement gap and average academic progress is imperative for understanding the observed differences between countries. We find no significant association between changes in national median and relative educational disadvantage over time. At the same time, as countries raise academic standards in reading and science, the gap in scores between the top and bottom achievers tends to narrow. So, there is no justified excuse for any educational system to ignore high levels of relative disadvantage. As we see on the examples of countries with high median performance (such as Estonia and Germany) as well as low average performers (such as Romania), the progress towards academic excellence can happen alongside targeted support of nations' low achievers or more equally shared improvement.

7.2 Absolute educational disadvantage

Failure to reach basic literacy competencies and skills among 15-year-olds provides a meaningful and internationally comparative scale of absolute educational disadvantage across the 39 industrialized countries in our study. We find stark differences between countries in terms of severe and overlapping low performance in all three subjects. Typically, this relates strongly to the overall level of academic achievement in a country. Despite the strong association, we find non-uniform ways in which educational systems approach extreme or cross-subject low performance likely to reflect specific educational policy goals. Some countries target performance of the most academically vulnerable groups (the case in Romania and Turkey), others provide for more equal support to all sub-groups of low performers without leaving the most academically vulnerable behind (the case of Bulgaria and Israel). Finally, there are countries that either focus on students who are at the margin of low performance with greater chances to climb above the threshold levels (the case of Japan), or allow the share of children with cross-subject low performance to increase to a greater extent than other low-performing sub-groups (the case of Sweden and Finland).

Popular macroeconomic indicators fall short of explaining variation in absolute educational disadvantage in the more homogeneous group of rich western economies. Therefore, there is nothing inevitable about letting children fall far behind their peers in educational performance. If Estonia, Latvia and Poland can achieve low shares of children falling behind minimum performance standards in maths, reading and science in spite of their modest resources, so could the richest economies, such as Luxembourg, Norway and the United States.

From the perspective of this paper the choice of addressing inequality of educational outcomes is that of redirecting limited resources to support the students with the greatest educational needs. This implies ensuring that as many children and young people as possible receive functional literacy skills above the minimum level. The key message of this paper is that the way in which the greater bottom-end equality of educational outcomes is achieved matters as much as the final outcome. As the analysis of relative disadvantage has shown, narrowing the gap might come with no direct benefit to the 'bottom group' but due to falling average standards. Therefore, policies of lifting up educationally disadvantaged children have to be integrated into the discussion of raising overall academic standards for all. Indeed, systematic and multifaceted policies to raise average achievement across the educational system are crucial, particularly for countries like Bulgaria, Chile, Mexico or Turkey. They would enable these educational systems to reduce unacceptably high levels of absolute educational disadvantage to levels comparable with other OECD countries.

The analysis has shown that 'a rising tide' is unlikely to 'lift all boats' (Kennedy 1963) in educational provision unless the most vulnerable are given the required learning support. Children who are 'hard to teach' and 'hard to reach' should not be left behind as chances for them to catch up with peers diminish with every academic year.

REFERENCES

Alderman, M. K. (2008). *Motivation for Achievement: Possibilities for Teaching and Learning.* New York: Routledge.

Atkinson, T. (2015). *Inequality. What can be done?* Cambridge, Massachusetts, London, England: Harvard University Press.

Berhanu, G. (2011). Inclusive Education in Sweden: Responses, Challenges and Prospects. *International Journal of Special Education*, *26 (2)*: 128-148.

Blanchenay, P., Burns, T., & Köster, F. (2014). Shifting Responsibilities - 20 years of Education Devolution in Sweden: A Governing Complex Education Systems Case Study. *OECD Working Papers,* No.104, OECD Publishing, http://dx.doi.org/10.1787/5jz2jg1rqrd7-en.

Blanden, J., & Machin, S. (2010). Intergenerational inequality in early years assessments. In K. Hansen, H. Joshi, & S. Dex (Eds.), *Children of the 21st Century–The First Five Years* (pp. 153-168). Bristol: The Policy Press.

Blanden, J. & McNally, S. (2014). *Reducing Inequality in Education and Skills: Implications for Economic Growth.* EENEE Analytical Report N21. www.eenee.de/dms/EENEE/Analytical_Reports/EENEE_AR21.pdf: European Expert Network on Economics of Education (EENEE).

Brighouse, H. (2014). Equality, prioritising the disadvantaged, and the new educational landscape. *Oxford Review of Education, 40(6):* 782-798.

Bulle, N. (2011). Comparing OECD educational models through the prism of PISA. *Comparative Education. 47(4)*: 503-521.

Cassen, R., & Kingdon, G. (2007). *Tackling low educational achievement.* York: Joseph Rowntree Foundation.

Causa, O. & Chapuis, C. (2010). 'Equity in student achievement in OECD countries'. *OECD Journal: Economic Studies, 5(1)*: 1-50.

Choi, A. & Jerrim, J. (2015). The use (and misuse) of PISA in guiding policy reform: the case of Spain? Department of Quantitative Social Science. Working Paper N.15-04. London: Institute of Education, UCL.

Cunado, J., & De Gracia, F. P. (2012). Does Education Affect Happiness? Evidence for Spain. *Social Indicators Research*, *108 (1)*, 185-196.

Currie, C., Currie, D., Menchini, L., Richardson, D., & Roberts, C. (2010). *Comparing Inequality in the well-being of children in economically advanced countries: A methodology.* Florence: Innocenti Working Paper, UNICEF.

Duncan, G.J & Murnane, R. Ed. (2011). *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances,* New York: Russel Sage Foundation.

Estonian Ministry of Education and Research (2015). Summary of PISA 2012 results for Estonia. Retrived from: http:www.innove.ee

Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort, *Economica*, 70(277): 73–97.

Ferreira, F.H.G. & Gignoux, J. (2011). The Measurement of educational inequality: Achievement and Opportunity. IZA Discussion Paper Series, No. 6161.

Freeman, R.B., Machin, S. & Viarengo, M. (2010). Variation educational outcomes and policies across countries and of schools within countries. *National Bureau of Economic Research Working Paper Series* No. 16293.

Hansen, K., & Joshi, H. (2007). *Millennium Cohort Study Second Survey: a user's guide to initial findings*. London: Centre for Longitudinal Studies.

Hermann, Z. & Horn, D. (2011). How inequality of opportunity and mean student performance are related? A Quantile regression approach using PISA data. Discussion papers. MT-DP- 2011/24, Budapest: Institute of Economics, Hungarian Academy of Sciences.

Jerrim, J., & Vignoles, A. (2013). Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(4): 887–906.

Jerrim, J. (2016). PISA 2012: How do results for the paper and computer tests compare? Working paper. UCL Institute of Education, Feb. 2016. Retrieved from: https://johnjerrim.files.wordpress.com/2013/07/working_paper_feb_2016.pdf

Kam, C.D. & Palmer, C.L. (2008). Reconsidering the effects of education on political participation. *The Journal of Politics, 70 (3)*: 612-631

Katz, I, Corlyon, J., La Placa, V. and Hunter, S. (2007). The relationship between parenting and poverty, York: Joseph Rowntree Foundation.

Kennedy, J. (1963). *Remarks in Heber Springs, Arkanzas, at the Dedication of Grers Ferry Dam.* October 3, 1963. Retrieved from The American Presidency Project: http://www.presidency.ucsb.edu/ws/index.php?pid=9455

Lareau, A. (2003). *Unequal Childhoods*, Berkeley: University of California Press.

Lindemann, K. (2013). The School Performance of the Russian-Speaking Minority in Linguistically Divided Educational Systems: A comparison of Estonia and Latvia. In M. Windzio, *Integration and Inequality in Educational Institutions* (pp. 45-71). New York, London: Springer.

McIntosh, S. (2004). The Impact of Vocational Qualifications on the Labour Market Outcomes of Low-Achieving School-Leavers, CEP Discussion Paper 621. London: LSE, Centre for Economic Performance

Micklewright, J. & Schnepf, S. (2006). Inequality of Learning in Industrialised Countries. IZA Discussion Paper Series, No. 2517.

Monseur, C. & Adams, R. (2009). Plausible values: How to deal with their limitations. *Journal of Applied Measurement, 10(3)*: 320-324.

OECD (2010). PISA 2009 Results: Overcoming Social background. Equity in Learning Opportunities and Outcomes. Volume II. http://dx.doi.org/10.1787/9789264091504-en: OECD.

OECD (2011). Divided We Stand: Why Inequality Keeps Rising, OECD Publishing, Paris, http://dx.doi.org/ 10.1787/9789264119536-en.

OECD (2012). Equity and Quality in Education: Supporting Disadvantaged Students and Schools. http://dx.doi.org/10.1787/9789264130852-en: OECD Publishing.

OECD (2013). Education at a Glance 2013. Country notes and key facts. Mexico;Turkey. http://www.oecd.org/edu/educationataglance2013-countrynotesandkeyfacttables.htm: OECD Publishing.

OECD (2014a). PISA 2012 results: What students know and can do. Students Performance in Mathematics, reading and science. (Volum I, Revised edition, February 2014). http://dx.doi.org/10/1787/97894201118.en: OECD Publishing.

OECD (2014b). PISA 2012 Technical report. http://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf: OECD Publishing.

OECD (2015). Improving schools in Sweden: OECD perspective. http://www.oecd.org/edu/school/Improving-Schools-in-Sweden.pdf: OECD Publishing.

Platt, L., Smith, K., Parsons, S., Connelly, R., Joshi, H., Rosenberg, R., et al. (2014). *Millennium Cohort Study: initial findings from the Age 11 survey*. Centre for Longitudinal Studies, Institute of Education.

Richardson, D. & Ali, N. (2014). An Evaluation of International Surveys of Children, DELSA/ELSA/WD/SEM(2013)2. www.oecd.org/els/workingpapers: OECD publishing.

Sahlgren, G. H. (2015). *Real Finnish Lessons. The true story of an education superpower.* London: Centre for Policy Studies.

Sanandaji, T. (2014). Sweden has an educational crises, but it wasn't caused by school choice. *National Review Online*, pp. http://www.nationalreview.com/agenda/383304/sweden-has-education-crisis-it-wasnt-caused-school-choice-tino-sanandaji.

Technical Advisory Group (TAG) of the EFA Steering Committee (2014). Towards Indicators for a post-2015 education framework. United Nations Educational, Scientific and Cultural Organisation. Accessed 15.06.2015, available at: http://www.uis.unesco.org/Education/Documents/towards-indicators-for-post-2015-education-framework.pdf

Thomas, V., Wang, Y., and Fan, X. (2001). Measuring educational inequality: Gini coefficients of education, Policy Research Working Paper 2525, Washington: The World Bank Group.

Townsend, P. (1979). *Poverty in the United Kingdom.* London: Allen Lane and Penguin Books.

UNDP (2013). *Humanity Divided: Confronting Inequality in Developing Countries.* United Nations Development Programme. New York: Bureau for Development Policy.

UNESCO (2012). *Beyond 2015: The Education We Want*. United Nations Educational, Scientific and Cultural Organisation. Accessed 07. 06.2015, available at: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/ED_new/Beyond2015_UNESCO-UNICEF-Flyer.pdf

UNICEF (2010). *The Children Left Behind: A league table of inequality in child well-beingin the world's rich countries.* Innocenti Report Card 9, Florence: UNICEF Innocenti Research Centre.

UNICEF (2002). *A League Table of Educational Disadvantage in Rich Nations. Innocenti Report Card Issue No. 4.* Florence: UNICEF, Innocenti Research Centre.

Uppsala Universitet (2014) School on a downward slope. *New Horizons online*, pp. http://www.uu.se/en/about-uu/new-horizons/article/?id=3516&area=15&typ=artikel&lang=en.

Waldfogel, J. (2013). Socio-economic inequality in childhood and beyond: an overview of challenges and findings from comparative analyses of cohort studies. *Longitudinal and Life Course Studies*, *4*(3), 268–275.

West, A. (2014). Academies in England and independent schools (fristående skolor) in Sweden: policy, privatisation, access and segregation. *Research Papers in Education, 29 (3)*, 330-350.

Willms, D. (2004). Variation in Literacy Skills among Canadian Provinces: Findings from the OECD Pisa. Education, skills and learning research papers. Ontario: Minister of Industry.

ANNEX

Table A.1 – *Country ranking based on three alternative measures of relative disadvantage*[21]

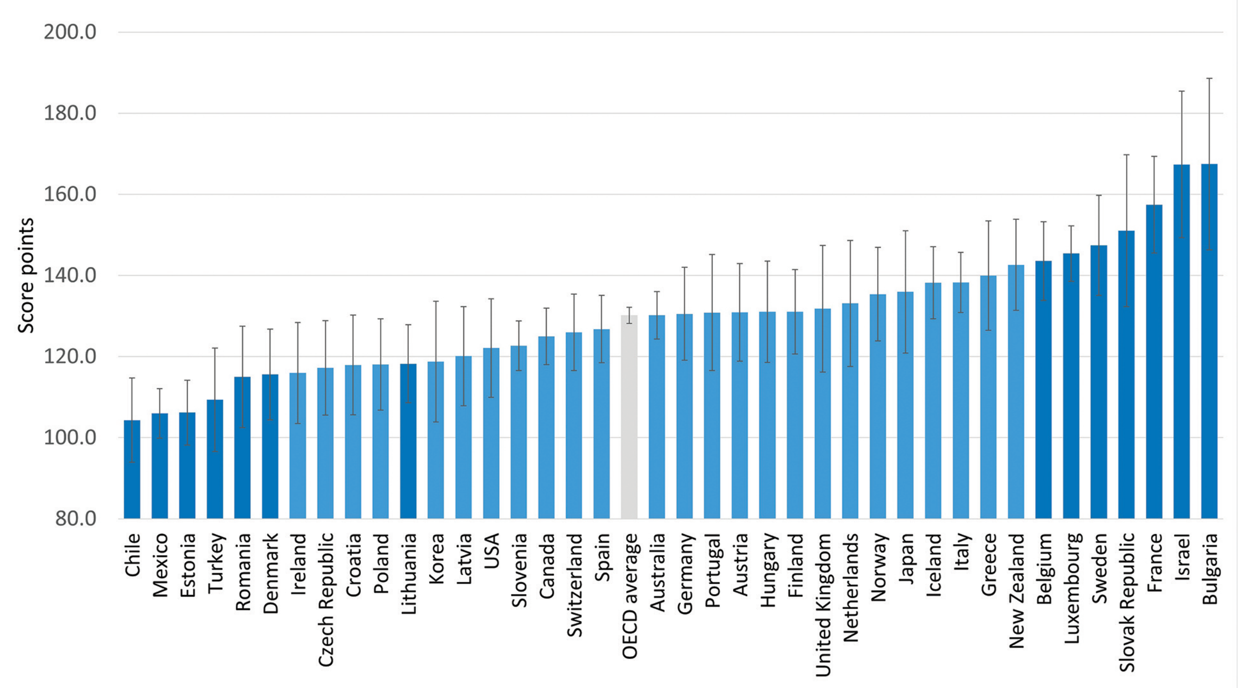
Country	SD
Mexico	**74.27**
Chile	**80.75**
Estonia	**80.90**
Romania	**81.34**
Latvia	**81.87**
Denmark	**82.10**
Ireland	**84.58**
Finland	**85.29**
Spain	**87.74**
Greece	**87.79**
Croatia	88.47
Canada	**88.86**
Lithuania	89.11
USA	89.86
Poland	90.37
Norway	90.48
Turkey	91.07
Netherlands	91.61
Slovenia	91.66
Sweden	91.75
OECD average	**91.86**
Iceland	91.94
Austria	92.48
Italy	92.78
Japan	93.52
Hungary	93.62
Bulgaria	93.91
Portugal	93.95
Switzerland	94.29
United Kingdom	94.52
Czech Republic	94.94
Luxembourg	**95.40**
Australia	**96.29**
Germany	**96.30**
France	**97.46**
Korea	**99.08**
New Zealand	**99.60**
Slovak Republic	**100.84**
Belgium	**102.26**
Israel	**104.91**

Country	50P-10P
Mexico	**91.37**
Chile	**93.91**
Romania	**94.53**
Turkey	**99.13**
Estonia	**101.29**
Latvia	**102.77**
Denmark	**107.10**
Croatia	108.05
USA	109.80
Finland	110.67
Ireland	111.61
Poland	112.35
Lithuania	113.72
Slovenia	113.97
Greece	114.47
Bulgaria	115.67
Spain	116.17
Hungary	116.26
Canada	116.40
Norway	117.25
Sweden	118.20
OECD average	**119.00**
Italy	119.86
Iceland	121.55
Australia	121.56
Austria	122.08
Czech Republic	122.96
Japan	123.66
United Kingdom	123.84
Portugal	125.08
Switzerland	125.80
New Zealand	127.33
Luxembourg	**128.00**
Slovak Republic	129.21
France	131.33
Germany	131.39
Korea	131.44
Netherlands	132.05
Belgium	**138.95**
Israel	**139.91**

Country	SD/mean
Estonia	**0.16**
Denmark	**0.16**
Finland	**0.16**
Latvia	**0.17**
Ireland	**0.17**
Canada	**0.17**
Japan	0.17
Poland	**0.17**
Netherlands	**0.18**
Switzerland	**0.18**
Korea	0.18
Mexico	**0.18**
Spain	**0.18**
Slovenia	0.18
Austria	0.18
Romania	0.18
Norway	0.18
OECD average	**0.19**
Lithuania	0.19
Iceland	0.19
USA	0.19
Germany	0.19
Croatia	0.19
Czech Republic	0.19
Australia	0.19
Chile	0.19
Italy	0.19
United Kingdom	0.19
Sweden	0.19
Portugal	**0.19**
Greece	**0.19**
Luxembourg	**0.19**
Hungary	0.20
France	**0.20**
Belgium	**0.20**
New Zealand	**0.20**
Turkey	**0.20**
Slovak Republic	**0.21**
Bulgaria	**0.21**
Israel	**0.22**

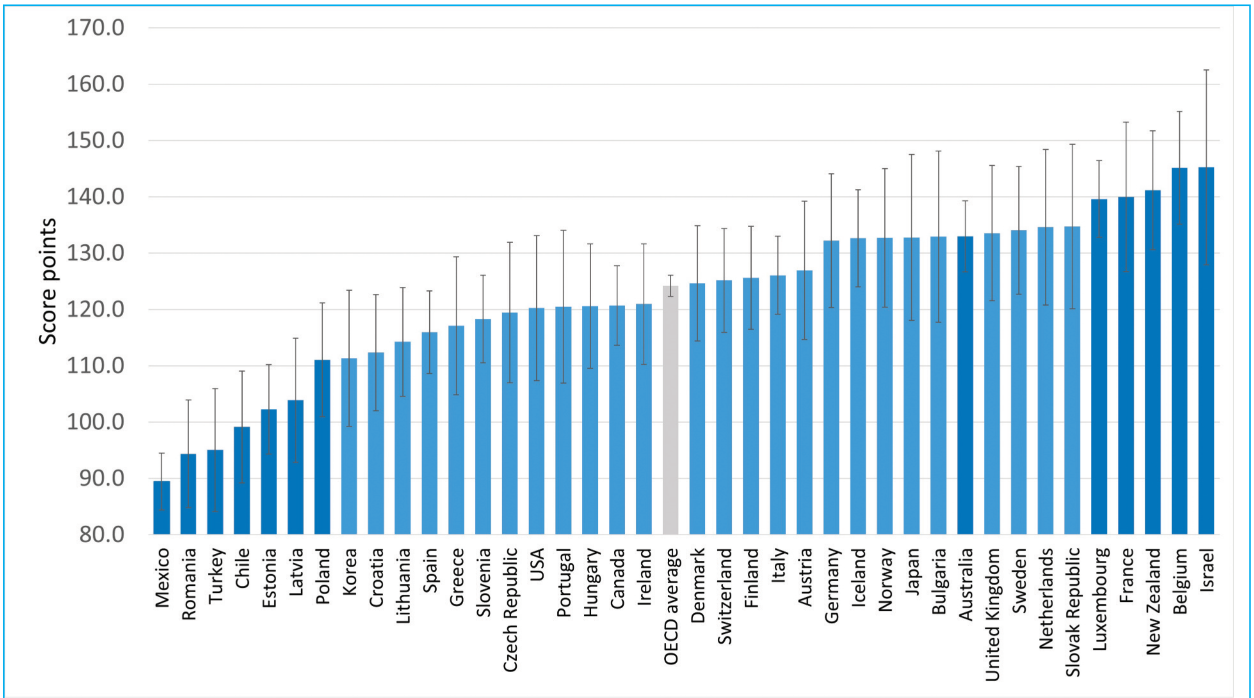
[21] Countries marked in bold are statistically different from the OECD unweighted average.

Figure A.1 – ***Score point achievement gap between children at the 10th percentile and children at the median (50th percentile) in reading in 2012***



Note: OECD unweighted average. Bars marked in the darkest colour show statistically different results from OECD average.
Source: PISA 2012.

Figure A.2 – ***Score point achievement gap between children at the 10th percentile and children at the median (50th percentile) in science in 2012***



Note: OECD unweighted average. Bars marked in the darkest colour show statistically different result from OECD average.
Source: PISA 2012.

Table A.2 – ***Score point change at the median, the 10th percentile and the gap between 50th and 10th percentile in mathematics***

	Change at the median 2003 - 2012		Change at the 10th percentile 2003 - 2012		Change in the bottom-end gap 2003 - 2012		Change at the median 2006 - 2012		Change at the 10th percentile 2006 - 2012		Change in the bottom-end gap 2006 - 2012	
	score	***sig***	***score***	***sig***	***score***	***sig***	***score***	***sig***	***score***	***sig***	***score***	***sig***
Australia	-24.1	***	-16.8	***	-7.3		-17.4	***	-25	***	7.2	
Austria	-0.2		-0.5		0.3		-4.5		11		-15.7	
Belgium	-20.3	***	-2.8		-17.5	**	-11.3	*	-3		-8.6	
Bulgaria	m		m		m		23.5	**	33	***	-9.1	
Canada	-15.6	***	-17.4	***	1.8		-10.9	**	-14	**	2.8	
Chile	m		m		m		8.8		21	***	-11.7	
Croatia	m		m		m		1.5		-1		2.6	
Czech Republic	-16.7	**	-14.5		-2.1		-9.6		1		-10.4	
Denmark	-15.6	***	-2.4		-13.1	*	-13.7	**	-11		-3.0	
Estonia	m		m		m		3.2		6		-3.3	
Finland	-24.0	***	-28.9	***	5.0		-30.4	***	-35	***	4.4	
France	-17.5	***	-23.3	**	5.8		-2.4		-3		0.7	
Germany	6.9		21.7	**	-14.8		11.0		10		1.2	
Greece	6.8		15.0	*	-8.1		-8.4		-2		-5.9	
Hungary	-15.7	**	-11.7		-4.1		-15.9	**	-19	**	3.1	
Iceland	-24.2	***	-23.7	***	-0.5		-12.9	***	-19	***	5.8	
Ireland	-0.9		-2.1		1.2		-0.1		-5		5.0	
Israel	m		m		m		26.5	***	24	**	2.2	
Italy	19.3	***	23.1	***	-3.8		23.4	***	24	***	-0.9	
Japan	-1.1		13.1		-14.2		12.7	*	11		1.8	
Korea	12.4	*	2.3		10.1		6.1		-1		7.4	
Latvia	5.4		15.9	*	-10.5		0.7		8		-7.5	
Lithuania	m		m		m		-8.9		-5		-3.8	*
Luxembourg	-3.8		-9.3	*	5.5		-0.7		-5		4.2	
Mexico	26.5	***	43.7	***	-17.2	**	5.3		21	***	-15.9	
Netherlands	-11.1		-18.9	*	7.8		-5.0		-16	*	10.6	
New Zealand	-26.5	***	-22.9	***	-3.6		-23.6	***	-29	***	5.5	
Norway	-4.9		-2.8		-2.1		0.1		0		0.1	
OECD average	-17.4	***	-3.2		-4.2		-14.9	**	-4.5		-0.4	
Poland	24.5	***	25.8	***	-1.3		19.6	***	18	***	1.5	
Portugal	21.0	***	11.2		9.8		20.2	**	15	*	5.5	
Romania	m		m		m		24.9	***	37	***	-12.3	
Slovak Republic	-17.5	**	-26.9	**	9.4		-13.4	*	-19	*	5.4	
Slovenia	m		m		m		-4.5		-6		1.7	
Spain	-1.1		1.2		-2.3		3.6		4		-0.2	
Sweden	-31.3	***	-27.0	***	-4.3		-24.2	***	-27	***	2.3	
Switzerland	3.8		12.0	*	-8.1		-0.2		7		-6.7	
Turkey	23.3	**	38.5	***	-15.2		22.8	***	23	***	0.2	
United Kingdom	-15.2	**	-17.0	*	1.8		0.2		-10		10.7	
United States	-6.1		11.1		-17.2	**	5.0		9		-4.4	

Source: PISA 2003, 2006, 2012. Standard errors (SE) are adjusted for link error: *** significant at 1%, ** significant at 2%, * significant at 5%.

Table A.3 – ***Score point change at the median, the 10th percentile and the gap between 50th and 10th percentile in reading***

	Change in the median 2000 - 2012		Change at the 10th percentile 2000 - 2012		Change in the bottom-end gap 2000 - 2012		Change in the median 2006 - 2012		Change at the 10th percentile 2006 - 2012		Change in the bottom-end gap 2006 - 2012	
	score	*sig*	*score*	*sig*	*score*	*sig*	*score*	*sig*	*score*	*sig*	*score*	*sig*
Australia	-18.6	*	-8.4		-10.1		-2.8		-2.8		0.0	
Austria	-19.1	*	-19.5	*	0.4		-3.1		17.3		-20.5	
Belgium	-9.2		13.6		-22.8	*	2.1		25.6	*	-23.6	*
Bulgaria	7.8		-20.9		28.7	*	38.9	**	24.7		14.2	
Canada	-11.5		-7.2		-4.4		-6.4		1.1		-7.5	
Chile	31.4	***	46.4	***	-15.0		0.9		29.4	**	-28.5	**
Croatia	m		m		0.0		5.1		10.1		-5.0	
Czech Republic	-4.3		3.3		-7.6		6.9		42.9	***	-36.0	***
Denmark	-3.0		19.3	*	-22.3	**	2.1		7.0		-4.9	
Estonia	m		m		0		13.4	*	23.2	**	-9.8	
Finland	-22.6	**	-29.7	***	7.13		-19.7	**	-41.4	***	21.7	**
France	3.4		-24.0	*	27.4	**	16.3	*	12.1		4.2	
Germany	18.6	*	41.2	***	-22.6	**	6.5		34.1	**	-27.6	*
Greece	7.6		6.0		1.6		16.3	*	24.3	*	-8.1	
Hungary	5.9		5.2		0.7		3.9		4.2		-0.3	
Iceland	-23.0	***	-31.5	***	8.5		-1.6		-4.0		2.4	
Ireland	-6.0		10.9		-16.9		4.4		14.9		-10.5	
Israel	37.1	**	26.0		11.2		51.6	***	49.6	***	2.0	
Italy	3.8		-9.0		12.8		19.4	**	34.6	***	-15.2	*
Japan	14.3		1.7		12.5		39.8	***	48.4	***	-8.7	
Korea	13.2		-8.9		22.1	*	-20.0	*	-16.2		-3.8	
Latvia	31.4	***	51.4	***	-20.0		12.0		13.8		-1.8	
Lithuania	m		m		m		7.7		20.7	**	-13.0	
Luxembourg	44.3	***	34.9	***	9.4		4.9		2.5		2.4	
Mexico	5.7		7.9		-2.1		10.5		34.5	***	-24.0	**
Netherlands	-20.6	*	-22.1		1.4		4.2		7.3		-3.1	
New Zealand	-21.2	**	-8.0		-13.2		-11.8		-7.2		-4.6	
Norway	-2.7		13.8		-16.5		18.1	*	28.8	**	-10.7	
OECD average	-1.3		3.0		-3.5		3.0		11.4		-7.6	
Poland	34.9	***	60.1	***	-25.1	*	9.6		30.2	***	-20.6	*
Portugal	17.7		24.8	*	-7.1		13.7		22.9	*	-9.3	
Romania	0.0		0.0		0		38.5	***	48.5	***	-10.0	
Slovak Republic	0.0		0.0		0		-0.8		-5.4		4.7	
Slovenia	0.0		0.0		0		-16.5	**	-15.1	*	-1.4	
Spain	-4.9		-11.3		6.5		26.0	***	23.8	**	2.15	
Sweden	-32.0	***	-47.6	***	15.6	*	-22.8	**	-35.8	***	13.0	
Switzerland	11.9		31.5	***	-19.6	*	8.7		14.8		-6.1	
Turkey	m		m		m		24.7	**	35.7	***	-11.0	
United Kingdom	-22.6	**	-18.0		-4.6		3.0		13.5		-10.4	
United States	-10.2		16.1		-26.3		0.0		m		m	

Source: PISA 2003, 2006, 2012. Standard errors (SE) are adjusted for link error: *** significant at 1%, ** significant at 2%, * significant at 5%.

Table A.4 – ***Score point change at the median, the 10th percentile and the gap between 50th and 10th percentile in science***

	Change at the median 2006- 2012		Change at the 10th percentile 2006 - 2012		Change in the bottom-end gap 2006 - 2012	
	score	*sig*	*score*	*sig*	*score*	*sig*
Australia	-6.5		-3.5		-2.9	
Austria	-6.4		4.9		-11.3	
Belgium	-4.9		-5.5		0.6	
Bulgaria	17.1	**	14.9		2.2	
Canada	-11.8		-2.3		-9.5	
Chile	8.4	**	20.3	**	-12.0	
Croatia	-1.0		-3.7		2.7	
Czech Republic	-2.8	**	7.5		-10.4	
Denmark	4.4	**	4.5		0.0	
Estonia	8.6		16.8	*	-8.2	
Finland	-15.9		-28.6	***	12.7	*
France	4.9	*	6.9		-2.0	
Germany	8.9	*	16.5		-7.6	
Greece	-7.9		-1.1		-6.8	
Hungary	-9.0		-12.2		3.2	
Iceland	-13.0	*	-16.5	**	3.5	
Ireland	15.2		18.6	*	-3.4	
Israel	21.2	**	18.1	*	3.1	
Italy	19.7		19.3	***	0.4	
Japan	14.8	**	24.2	*	-9.4	
Korea	15.8	***	27.4	**	-11.6	
Latvia	12.4		20.1	**	-7.6	
Lithuania	8.2	*	13.4	*	-5.3	
Luxembourg	3.8	**	-3.4		7.1	
Mexico	7.2	***	19.3	**	-12.1	*
Netherlands	-2.3		-1.6		-0.7	
New Zealand	-15.7		-12.0		-3.6	
Norway	9.9		0.5		9.5	
OECD average	1.3		4.9		-3.6	
Poland	28.4	***	33.9	***	-5.5	
Portugal	16.0	*	14.2	*	1.7	
Romania	18.8	*	26.3	***	-7.4	
Slovak Republic	-14.5		-28.3	***	13.7	
Slovenia	-3.9		6.2		-10.1	
Spain	8.6		13.8	*	-5.3	
Sweden	-17.2		-27.2	***	10.0	
Switzerland	2.8		15.8	*	-12.9	
Turkey	42.5		38.3	***	4.2	
United Kingdom	0.2		8.6		-8.4	
United States	9.3		27.9	***	-18.5	

Source: PISA 2003, 2006, 2012. Standard errors (SE) are adjusted for link error: *** significant at 1%, ** significant at 2%, * significant at 5%.

Table A.5 - ***Share of children below level 2 in 3 subjects (the overlap group).***

	2008	2009	2012		
	Overlap group (bellow level 2 in 3 subjects)	Overlap group (below level 2 in 3 subjects)	Overlap group (below level 2 in 3 subjects)	Change 2006-2012	Change 2009-2012
Australia	7.14	8.48	9.10	1.95	0.62
Austria	10.73	15.14	10.71	-0.02	-4.43
Belgium	10.87	11.74	11.53	0.66	-0.21
Bulgaria	35.06	29.99	28.57	-6.49	-1.42
Canada	5.20	5.16	6.16	0.96	0.99
Chile	25.26	22.88	24.57	-0.69	1.69
Croatia	12.21	13.79	11.71	-0.50	-2.09
Czech Republic	10.63	11.75	8.91	-1.72	-2.84
Denmark	8.21	8.84	9.33	1.13	0.49
Estonia	5.00	5.25	3.25	-1.75	-2.00
Finland	1.88	3.14	5.35	3.47	2.21
France	13.07	13.51	12.66	-0.41	-0.85
Germany	10.96	10.22	8.76	-2.20	-1.46
Greece	15.46	13.80	15.75	0.28	1.95
Hungary	10.58	11.01	13.10	2.52	2.09
Iceland	10.49	10.03	13.56	3.07	3.53
Ireland	7.53	10.65	6.76	-0.77	-3.89
Israel	25.90	21.25	18.50	-7.39	-2.74
Italy	15.19	12.74	11.87	-3.32	-0.87
Japan	7.69	7.22	5.52	-2.17	-1.70
Korea	3.88	3.29	4.35	0.47	1.06
Latvia	10.32	8.92	8.29	-2.03	-0.63
Lithuania	13.45	12.62	12.07	-1.39	-0.55
Luxembourg	14.41	15.66	14.40	-0.01	-1.26
Mexico	35.20	31.82	30.99	-4.20	-0.82
Netherlands	7.36	6.82	8.62	1.26	1.80
New Zealand	7.81	8.50	11.14	3.33	2.64
Norway	12.17	9.04	11.00	-1.17	1.96
OECD average	11.67	11.47	11.61	-0.06	0.13
Poland	9.42	8.57	5.71	-3.72	-2.86
Portugal	16.36	10.39	12.56	-3.80	2.18
Romania	35.89	29.77	24.02	-11.87	-5.75
Slovak Republic	13.07	11.52	18.77	5.71	7.25
Slovenia	7.95	10.37	9.93	1.98	-0.44
Spain	13.05	11.47	10.37	-2.68	-1.09
Sweden	8.93	11.06	15.00	6.06	3.93
Switzerland	8.99	7.72	7.46	-1.53	-0.26
Turkey	25.38	18.62	15.56	-9.82	-3.06
United Kingdom	10.87	10.45	11.19	0.31	0.73
United States	0.00	12.03	12.22	0.00	0.19

Source: PISA 2012