Original Research Article

Dividing Students, Dividing Outcomes: Age at Selection, Ability Grouping, and Educational Equity in High-Performing Education Systems

Sangjin Lee

St. Mary's International School, 1-6-19 Seta, Setagaya-ku, 158-8868, Japan

ABSTRACT

This study examines the association of age at selection, ability grouping, and mandatory standardized testing with socio-economic fairness (SEF) in mathematics in Switzerland and Estonia. Using the Organization for Economic Co-operation and Development Program for International Student Assessment (OECD PISA) 2022, this research selects Switzerland and Estonia as contrasting cases due to their similar high performance in mathematics but different levels of stratification and accountability systems. The Pearson correlation coefficient analysis of SEF and the age at selection shows a moderate-to-strong positive correlation ($r = 0.4995$, $n = 21$), whereas there is a weak negative correlation between SEF and ability grouping ($r = -0.1125$, $n = 23$). Descriptive comparison between Switzerland and Estonia provides no conclusive evidence to support that higher stratification is associated with a lower sense of belonging and self-confidence. In addition, a moderate positive correlation is found between SEF and mandatory standardized testing ($r = 0.4799$, $n = 17$), although this correlation should not be interpreted as causal. Overall, this paper's findings suggest that earlier age at selection is associated more consistently with lower SEF in mathematics in high-performing education systems. The association between SEF in mathematics and mandatory standardized testing is positive but only suggestive, while the association between SEF in mathematics and ability grouping remains inconclusive.

Keywords: Educational equity; socio-economic fairness; ability grouping; age at selection; PISA 2022; Switzerland; Estonia; Mathematics achievement

INTRODUCTION

Education inequality remains a persistent issue even within high-achieving education systems. While many countries aim to achieve higher performance, the academic performance gap between economically advantaged and economically disadvantaged students persists significantly and is often overlooked. According to the Organization for Economic Co-operation and Development (OECD), on average across OECD countries, "economically advantaged students scored 93 points more in mathematics than disadvantaged students," and this trend is projected to increase (1). This trend showcases that the distribution of educational outcomes is strongly influenced by students' socio-economic status (SES).

Beyond socio-economic backgrounds, education policies and systems may also affect how equitably academic opportunities and outcomes are distributed across social groups. Stratification is defined as the

Corresponding author: Sangjin Lee, E-mail: Sjlee121208@gmail.com.

Accepted July 31, 2026
https://doi.org/10.70251/HYJR2348.44616629

separation of students into separate groups, tracks or streams for instructional purposes (2). In this research, it is used as an umbrella term for all policies that separate and differentiate students within class. Specifically, since the Coleman report, the associations of the age at selection and ability grouping with educational equity have been widely examined in previous studies (3). The age at selection is the age at which students are first separated into different educational programs, and ability grouping is defined as grouping students by ability on academic performance (3). These factors are considered key determinants of socio-economic fairness (SEF), defined as the percentage of variation in how much the disparity of scores between economically advantaged students and disadvantaged students is independent of SES (4, 5). Due to the multifaceted factors affecting an education system's SEF, isolating a specific policy and its association with SEF has been challenging. Thus, it has been an important research objective to empirically identify these policies that affect the SEF in education. Among these policies are mandatory standardized testing, which has been found to increase academic scores and may be associated with a smaller academic performance gap between disadvantaged and advantaged students (6).

There have been many previous studies regarding SEF and education systems. Nevertheless, disputes prevail. First, although existing research has examined the effect of ability grouping and earlier age at selection, the results mainly revolve around academic achievements rather than educational inequality (6, 7). Furthermore, the results regarding educational inequality are clouded by the fact that academic achievements were so strongly correlated with the students' SES to reliably conclude that ability grouping and the age at selection were the reasons behind the gap between disadvantaged students and advantaged students (7). Second, while past studies have found that mandatory standardized testing is effective in increasing academic scores, its association with educational inequality remains unclear (6, 9, 10). Furthermore, discrepancies in the results exist between different academic subjects and case studies (11). Thus, to address the discrepancies in previous studies, this study explores how ability grouping, earlier age at selection, and mandatory standardized testing are associated with educational equity using OECD PISA 2022's data on SEF in mathematics.

Switzerland and Estonia offer a useful most-similar-systems comparison to examine the association of these policies with SEF in mathematics: both are considered high-performing in the Organization for Economic Co-operation and Development's (OECD) Programme for International Student Assessment (PISA) 2022, yet they differ in their educational policies regarding age at selection and ability grouping (12). The comparison allows the study to examine SEF and its associations, while the regression analysis helps contextualize whether the associations found between Switzerland and Estonia appear beyond the two education systems in similar high-performing education systems. The study therefore asks: To what extent are earlier ages at selection and ability grouping associated with lower socio-economic fairness in mathematics in similarly high-performing education systems, as illustrated by Switzerland and Estonia in PISA 2022? The objective of this study is to examine whether ability grouping, age at selection, and mandatory standardized testing are predominantly positively, negatively, or conditionally associated with the education system's SEF in mathematics among high-performing education systems, and also to assess the association of these policies with self-concept.

This research tests two hypotheses and one exploratory research question:

- (1) Higher stratification, including higher ability grouping and earlier age at selection, is associated with lower SEF in mathematics
- (2) Exploratory research question: is higher stratification associated with lower disadvantaged students' self-confidence and sense of belonging?
- (3) More prevalent mandatory standardized testing is associated with higher SEF in mathematics

LITERATURE REVIEW

Self-concept in Education Inequality

Self-concept, or the sense of who they are that individuals derive from their membership in social groups, is integral to understanding students' academic behaviors (8). Self-concept includes various components, including the sense of belonging and self-confidence (13, 14). It is well established in previous studies that a strong establishment of self-concept leads to proactive and positive attitudes of students, and naturally higher academic achievements (13). Conversely, a lower self-concept is associated with a lower rate of increase in school grades and lower school achievements. According to Jetten and Peters, disadvantaged students are more likely to experience a lower self-concept, including a diminished sense of belonging, leading to lower overall morale and motivation needed to achieve academic advancement (8). Consequently, this results in identity incompatibility, defined as the conflict between identity-related aspects, due to students believing that they are “not the type of person who does well in school” (8, 15). As such, students grouped in lower ability classes may be affected by identity incompatibility and therefore perform worse in academics (8).

Research from the University of London conducted by Harrold, Fett, and Haenschel in 2024 empirically illustrates that disadvantaged students have a lower self-concept than advantaged students in secondary education, and increased stratification of subject levels by ability further widens the gap (15). Nonetheless, self-confidence was highest in settings with moderate levels of ability grouping (13).

The research further establishes a chain between self-concept and educational outcomes. Class-based stereotypes are reinforced by social and cultural factors, which then translate into suppressed educational outcomes due to identity incompatibility; such suppressed education outcomes then translate back into a lower self-concept, forming a continuous cycle (8). Nevertheless, this cycle can be broken through positive representations, such as setting a role model that students can follow (8). Research from Janet Schofield of the University of Pittsburgh finds that students’ math achievements are strongly correlated with their peers’ ability; furthermore, Schofield states that high-achieving students from low SES were most affected by their peers (7). In this context, the peers act as role models where students, especially those from lower SES, strive to achieve equally higher performance in academics. Unfortunately, the interpretation is clouded by the fact that students’ SES and academic achievement were so strongly correlated that the researchers omitted mean SES as a control variable (7).

Ability Grouping and Education Inequality

Ability grouping is a component of stratification that aims to achieve higher efficiency for all students by separating those with similar skills in classes tailored to their abilities (11). Research by Blanco-Varela, Amoedo, and Sanchez-Carreira shows that in Spain, across subjects, disadvantaged students (as shown by high vulnerability) demonstrated lower performance compared to advantaged students (as shown by low vulnerability) within ability grouping (11). At the same time, the knowledge area regarding reading showed the opposite trend, with advantaged students showing lower performance compared to disadvantaged students. Likewise, there is a discrepancy in that moderate levels of socio-educational vulnerability are associated with the least performance variation across global subjects, including math, science, and reading.

Rather than employing ability grouping, the research points out multiple benefits of heterogeneous or mixed-ability grouping, including facilitation of interaction of diverse abilities, reducing stigma associated with low achievers, and setting high achievers as role models to their peers (11). However, considering that this research was conducted in the secondary education system of Spain, it may carry limitations in the application to Switzerland and Estonia. Furthermore, although many studies assert that mixed-ability grouping improves performance among disadvantaged students as it weakens the social identity related to them, research by Fleischmann, Trautwein, Hübman, and Nagengast suggests that mixed-ability grouping, or detracking as they call it, conversely lowers the performance of disadvantaged students due to demotivation (16). Particularly, the study finds that detracking comes at the cost of lowering self-concept among low achieving and disadvantaged students individually, ultimately indicating lower academic performance (16)

Early Selection and Socioeconomic Inequality

The logic behind earlier age at selection and its effect on education equity is that the earlier children are separated by ability or merit, the stronger the influence of Economic, Social, and Cultural Status (ESCS), measured as the average of indices of parental education attainment, parental occupational status on the “International Socio-Economic Index,” and household possessions (17). In turn, more homogenous groups of children will go to the same schools or classes, reinforcing the education inequality and increasing the gap between disadvantaged and advantaged students (18). Earlier selection does not only produce negative results, though. The formation of homogeneous classes by the education system may benefit students of high ability due to the positive peer-group effect, as found before in Schofield’s research. Nonetheless, with lower-status families already considering education selection as a “hurdle,” disadvantaged students will over time be concentrated, resulting in a low positive peer-group effect within the class. Furthermore, through such a selection system, students are placed into different education tracks that lead to jobs with different labor market value (7). Hence, not only is the earlier age at selection system widening education inequality, but it can also become a key determinant of future wages.

While multiple other studies also find that greater stratification tends to exacerbate inequality of educational outcomes (IEO) within general secondary education, a significant limitation persists in that intensified academic competition within unstratified compulsory education likewise leads to wider IEO (19). As such, although stratification is strongly associated with educational inequality, it is contested whether the IEO directly results from stratification or from external factors, such as academic competition embedded since compulsory education (9). Thus, while it is not necessarily always true, it is possible that the more stratified a system is, the greater the socioeconomic inequality in educational outcomes.

Standardization

Standardization is a broad concept that encompasses a variety of aspects of schooling (8). Due to its broad and vague nature, it often uses its sub-concepts to define itself, which includes accountability and centralization. Accountability is defined as a system of standardized incentives for schools; it is an umbrella term to refer to any policies that may be associated with higher accountability (9). The relationship between accountability and standardization of education is that the more homogenous the accountability system is, the more standardized education is. One of the most common forms of accountability is nationally standardized testing. Many studies have shown the effectiveness of standardized testing in decreasing the performance gap between disadvantaged and advantaged students, including Bishop and Mane's, which also found that voluntary or non-standardized exit exams are not effective (6).

Centralization is defined by the strength of central regulation on school processes (9). Similar to accountability, the more the central government is involved in schools' decision-making, the more standardized the education system is. With weaker centralization, or stronger decentralization, school autonomy is stronger, meaning that there is more decision-making by schools. This can include the ability to choose curricula, teacher tutoring policies, and peer-to-peer tutoring policies (9). Although there are results that show the correlation between centralization and decreased education inequality, the data are limited to the US only, making the application to either Switzerland or Estonia unsubstantiated (10). Moreover, the majority of the research explores the association between accountability and centralization on education outcome and performance, leaving the effect on education inequality unanswered.

METHODS AND MATERIALS

This study builds on previous studies exploring the association of age at selection, ability grouping, and mandatory standardized testing with SEF. More specifically, this research focuses on the Swiss Confederation (Switzerland) and the Republic of Estonia (Estonia). This is not to say that Switzerland and Estonia are the only possible comparison, but because Switzerland and Estonia provide a most-similar-systems comparison as they are both academically high-performing, with strong educational systems and resources as indicated by OECD PISA 2022 data (20). The geographical and geopolitical proximity of Estonia and Switzerland allows for an easier comparison, minimizing external factors such as cultural differences. Furthermore, both Switzerland and Estonia are high-income OECD countries, which would help to reduce the extent to which differences in national income affect the comparison (21, 22). Meanwhile, they differ in school-level support and accountability systems, including the use of mandatory standardized testing, internal/self-evaluations, teacher mentoring, and peer-to-peer tutoring. As such, Switzerland and Estonia's similarities allow for a more focused comparison, isolating differences in age at selection and ability grouping as key variables for variations in SEF.

Data for education systems globally, including Switzerland and Estonia, were obtained from the OECD PISA, which is publicly accessible. The data, OECD PISA 2022 Results Volume 1 and Volume 2, were downloaded from the OECD website directly on March 15th, 2026 (12). To assess the association of different policies and SEF among high-performing education systems, this study exclusively uses values of education systems whose performance is statistically significantly above the OECD average at the 5% level ($p < 0.05$) as indicated by OECD PISA 2022 (12). OECD's significance criteria were adopted rather than a numerical cutoff because small differences with the OECD average may fall within sampling and measurement uncertainty and may not represent statistically meaningful differences in performance. From 81 education systems that participated in the OECD PISA 2022, 23 met the OECD significance criterion. The selected education systems include Switzerland, Estonia, New Zealand, Slovenia, Korea, Japan, Australia, Austria, United Kingdom, Latvia, Sweden, Finland, Czech Republic, Netherlands, Belgium, Ireland, Canada, Singapore, Denmark, Macao (China), Chinese Taipei, Hong Kong (China), and Poland (n = 23); not in particular order. Education systems with unavailable values were excluded only from the relevant analysis, resulting in n = 23 for ability grouping, n = 21 for age at selection, and n = 17 for mandatory standardized testing. The OECD average was included in tables only as a descriptive benchmark and was excluded from all analyses because it is not an independent education-system-level observation.

Education system observations were contained in this dataset, including student and school faculty surveys, mean performance for subjects (math, science, reading), SEF index, trend across ESCS quartiles for performance in different subjects, performance between genders, education inequality between immigrants and non-immigrants, and rate of grade repetition. Overall, the SEF in mathematics, mean performance for mathematics in the education system, first age at selection, and percentage of students attending a school that groups students based on different abilities were used as main variables in this study. No composite index was constructed to assess self-concept. Instead, sense of belonging, self-confidence, and isolation index among disadvantaged and advantaged students were analyzed and reported separately to descriptively assess self-concept.

All variables and shortened terms utilized in this study are mapped directly to their OECD PISA 2022:

Economic, Social, and Cultural Status (ESCS): refers to PISA's index of student socio-economic status, based on parental education, parental occupational status, and household possessions. In this study, ESCS was not analyzed as a separate indicator, but it was used to measure students' socio-economic status within the OECD PISA 2022 data, such as defining "advantaged students" and "disadvantaged students." ESCS was then used to calculate the SEF in mathematics.

Socio-economic fairness (SEF): written as "Socio-economic fairness in mathematics," this system-level indicator is measured by the percentage of variation in student performance that is not accounted for by differences in student socio-economic status. Higher percentages indicate higher levels of fairness by student socio-economic status. It is mathematically calculated as 100 minus the percentage of total variance explained by the ESCS index. SEF was derived from OECD, PISA 2022 Database, Annex B1, Chapter 1; and Volume I, Annex B1(12).

Ability grouping: derived from OECD, PISA 2022 Database, Annex B1, Chapter 4 and Table B3.1.4, it refers to the "percentage of students who attended a school where students are grouped by ability into different classes for all subjects." This functions as a stratification policy (18).

Age at selection: derived from OECD, PISA 2022 Database, Annex B1, Chapter 4 and Table B3.1.4, it refers to the "First age at selection in the education system." This system-level metric represents the official age at which an education system first separates students into distinct academic or vocational courses. This functions as a stratification policy, where earlier age at selection indicates earlier stratification (12).

Mandatory standardized testing: derived from OECD, PISA 2022 Database, Annex B1, Chapter 6, Table II.B1.6.39, it refers to the "percentage of students in schools where students are assessed through mandatory standardized tests at least once a year." It was treated as an accountability indicator within the broader concept of standardization (12).

Internal/self-evaluations: derived from OECD, PISA 2022 Database, Annex B1, Chapter 6, Table II.B1.6.59, it refers to the "percentage of students in schools where internal evaluations or self-evaluations are in place" (12). Internal/self-evaluations were treated as an accountability indicator within the broader concept of standardization. The values were used to descriptively compare Switzerland and Estonia.

Teacher mentoring: derived from OECD, PISA 2022 Database, Annex B1, Chapter 6, Table II.B1.6.66, it refers to the "percentage of students in schools where teacher mentoring is in place" (18). Teacher mentoring was treated as a school-level support indicator and used to descriptively compare Switzerland and Estonia.

Peer-to-peer tutoring: derived from OECD, PISA 2022 Database, Annex B1, Chapter 5, Table II.B1.5.85, it refers to the "percentage of students in schools where peer-to-peer tutoring is provided" (12). Peer-to-peer tutoring was treated as a school-level support indicator and used to descriptively compare Switzerland and Estonia.

Sense of belonging written as "Index of sense of belonging," it was derived from OECD, PISA 2022 Database, Annex B1, Chapter 1, Table II.B1.1.1 (12). The index values are extracted through student responses that indicate their level of belonging in school by rating their agreement with six distinct statements on a 4-point Likert scale (*Strongly disagree*, *Disagree*, *Agree*, *Strongly agree*): *'I feel like an outsider (or left out of things) at school'*; *'I make friends easily at school'*; *'I feel like I belong at school'*; *'I feel awkward and out of place in my school'*; *'Other students seem to like me'*; and *'I feel lonely at school'*. The value ranges from -1 to 1, where a negative value corresponds to below the OECD average of the sense of belonging at 0, while a positive value corresponds to a higher sense of belonging than the OECD average. Sense of belonging was used as a component of indicating self-concept to descriptively assess exploratory research question 2.

Self-confidence: written as "percentage of students who feel confident or very confident that they can motivate themselves to do schoolwork," the values were derived from OECD, PISA 2022 Database, Annex B1, Chapter 2, Table II.B1.2.5. The percentages were calculated through students' self-responses on a 4-point confidence scale: '*Not at all confident,*' '*Not very confident*,' '*Confident*,' and '*Very confident.*' Self-confidence was used as a component of indicating self-concept to descriptively assess exploratory research question 2 (12).

Isolation index: derived from OECD, PISA 2022 Database, Annex B1, Chapter 4 and Table II.B1.4.17, the isolation index measures the extent to which certain types of students (e.g. disadvantaged students) are isolated from all other types of students or from a specific group of students (e.g. advantaged students) based on the schools they attend. It ranges from 0 to 1, where 0 corresponds to full exposure (no segregation) and 1 to full isolation/segregation. The values used for this study were extracted under the isolation index for "Disadvantaged students from all other students" and "Advantaged students from all other students" (12). The isolation index was used as a contextual measure that may relate to students' sense of belonging.

Changes in mathematics performance: derived from OECD, PISA 2022 Database, Annex B1, Chapter 1, Table II.B1.2.46, "Change in mathematics performance" (12). It was measured by subtracting the PISA mathematics performance mean score from PISA 2022 mathematics performance mean score. The values were used to descriptively assess exploratory research question 2, specifically that higher stratification may result in lower academic performance in mathematics.

Although self-reported measures are often subject to cognitive biases, their use in this study is both appropriate and necessary. Constructs such as self-confidence and sense of belonging are inherently subjective and cannot be measured objectively. Hence, self-reported indicators are widely utilized to address these internal states. Furthermore, considering that the OECD PISA 2022 survey employed a standardized questionnaire for all education systems, any self-reporting biases were likely consistent throughout, preserving the validity of the data.

For all Tables 1, 2, 3, and 4, values numerically above the OECD average were highlighted in blue and values numerically below the OECD average were highlighted in gray, regardless of statistical margins. The OECD average and unavailable values noted in N/A were not highlighted in any color. Color-coding is only descriptive and does not indicate statistical significance or causal effect. Statistical interpretation is based on the Pearson correlation coefficient, linear regression model, P-value, coefficient of determination (R^2), 95% confidence interval, standard error of slope, and diagnostic assessment. SEF in Mathematics was measured by the percentage of variation in student performance that is not accounted for by differences in student SES. Higher percentages indicate higher levels of fairness by student SES. Self-confidence in education was measured by the percentage of students who reported that they feel confident or very confident that they can motivate themselves to do schoolwork.

OECD PISA 2022 did not report standard errors or confidence intervals for every system-level indicator used in this study's analysis. Hence, OECD uncertainty estimates were not directly incorporated into the regression models. The unit of analysis was the education system, not the individual student. Pearson correlation coefficients were then calculated to examine system-level correlations between SEF and ability grouping, age at selection, and mandatory standardized testing.

To first analyze the correlation between SEF in mathematics and age at selection and ability grouping, the Pearson correlation coefficient (r) was calculated using a TI-84 calculator. The formula can be represented as:

$$r_{(SEF,\ Age\ at\ Selection)} = \frac{\Sigma\left(X_1 - \underline{X}_1\right)\left(Y - \underline{Y}\right)}{\sqrt{\Sigma\left(X_1 - \underline{X}_1\right)^2 \cdot \Sigma\left(Y - \underline{Y}\right)^2}}$$

$$r_{(SEF,\ Ability\ Grouping)} = \frac{\Sigma\left(X_2 - \underline{X}_2\right)\left(Y - \underline{Y}\right)}{\sqrt{\Sigma\left(X_2 - \underline{X}_2\right)^2 \cdot \Sigma\left(Y - \underline{Y}\right)^2}}$$

Calculating the value, the r value for (SEF, ability grouping) comes out at -0.1125, indicating a weak negative correlation, while the r value for (SEF, Age at selection) comes out at 0.4995, indicating a stronger positive correlation. A correlation closer to 1 indicates a stronger relationship, with a positive value indicating that both X and Y will increase; a negative value indicates that an increase in the X value correlates with a decrease in the Y value. Using the linear regression, the fitted regression model for $r_{(SEF,\ age\ at\ Selection\ in\ the\ education\ system)}$ was calculated as $y = 68.33 + 1.188x$

$r_{(SEF,\ ability\ grouping)}$ was calculated as $y = 86.09 - 0.05712x$

As Table 2 shows, Switzerland and Estonia notably differ from the index of sense of belonging and self-confidence in education. Nonetheless, the isolation index differences of each disadvantaged and advantaged student were small and interpreted descriptively.

Table 1. *Mean mathematics performance, socio-economic fairness (SEF) in mathematics, performance gap between disadvantaged and advantaged students, first age at selection in the education system, and percentage of students attending schools where ability grouping is present among OECD-defined high-performing PISA 2022 systems. The table includes n = 23 education systems, with the OECD average shown only as a benchmark. For analyses, OECD average was excluded. Missing values marked as N/A are as follows: Canada and Australia for "First age at selection in the education system." Blue = above OECD average; gray = below OECD average; colors are descriptive only.*

Education System	Mean Mathematics Score	Difference between advantaged and disadvantaged students in mathematics	Socio-economic fairness (SEF) in Mathematics	First age at selection in the education system	Percentage of students who attended schools where students are grouped by ability into different classes for all subjects
OECD Average	472	93	84.5%	14.3	6.70%
Switzerland	508	117	79.2%	12	26.1%
Estonia	510	81	86.6%	16	6.30%
Canada	497	76	89.8%	N/A	8.22%
Netherlands	493	106	84.9%	12	37.2%
Korea	527	97	87.4%	15	8.33%
Singapore	575	112	83.0%	12	7.31%
Ireland	492	74	87.0%	15	0.63%
Belgium	489	117	78.2%	15	10.1%
Denmark	489	74	87.8%	16	1.40%
United Kingdom	489	86	89.0%	16	5.04%
Poland	489	96	83.7%	15	2.98%
Japan	536	81	88.1%	15	6.20%
Austria	487	106	80.6%	10	3.50%
Australia	487	101	85.4%	N/A	2.70%
Czech Republic	487	116	78.0%	11	2.92%
Slovenia	485	91	84.3%	15	10.0%
Finland	484	83	87.6%	16	0.92%
Latvia	483	75	86.8%	16	6.58%
Sweden	482	99	85.0%	16	0.00%
New Zealand	479	102	84.2%	16	1.43%
Macao (China)	552	55	95.0%	15	6.31%
Chinese Taipei	547	119	84.3%	15	6.31%
Hong Kong (China)	540	65	94.2%	14	13.2%

Table 2. *Student belonging, isolation, confidence, and mathematics performance-change indicators for Switzerland and Estonia, with the OECD average included as a descriptive benchmark. Isolation indices range from 0 to 1, with higher value indicating higher isolation. Index of Sense of Belonging ranges from -1 to 1, where a positive value indicates a higher sense of belonging than the OECD average of 0. The table includes n = 2 education systems, with the OECD average shown only as a benchmark. Blue = above OECD average; gray = below OECD average; colors are descriptive only.*

Education System	Self-confidence in education	Index of Sense of Belonging	Isolation Index of:		Change in mathematics performance for:	
			Disadvantaged students from all other students	Advantaged students from all other students	Disadvantaged students	Advantaged students
OECD Average	58.1%	-0.02	0.18	0.19	-17	-10
Switzerland	64.8%	0.36	0.15	0.20	-23	-6
Estonia	56.3%	-0.14	0.17	0.18	-15	-2

Table 3. *Standardization indicators among OECD-defined high-performing education systems in PISA 2022. Values are percentages of students attending schools where each indicator is reported. The table includes n = 23 education systems, with the OECD average shown only as a benchmark and excluded from all analysis; the standardized-testing analysis used n = 17 education systems after excluding N/A values. Missing values marked as N/A are as follows: Netherlands, Ireland, Japan, Australia, New Zealand, and Hong Kong (China) for "Percentage of students in schools where: Students are assessed through mandatory standardized tests at least once a year." Blue = above OECD average; gray = below OECD average; colors are descriptive only.*

Education Systems	Percentage of students in schools where:			
	Teacher Mentoring is in place	Peer-to-peer tutoring is provided	Students are assessed through mandatory standardized tests at least once a year	Internal evaluations or self-evaluations are in place
OECD Average	81.9%	51.3%	72.5%	95.3%
Switzerland	83.3%	28.5%	65.6%	84.8%
Estonia	96.0%	53.3%	98.1%	99.7%
Canada	90.9%	71.1%	83.0%	83.3%
Netherlands	94.1%	43.2%	N/A	96.9%
Korea	98.0%	78.8%	75.6%	99.6%
Singapore	100%	65.8%	97.6%	99.1%
Ireland	95.7%	19.1%	N/A	100%
Belgium	92.1%	27.8%	33.6%	89.8%
Denmark	88.3%	20.7%	80.2%	88.6%
United Kingdom	96.8%	53.7%	95.4%	100%
Poland	94.8%	83.7%	55.6%	89.6%
Japan	87.5%	47.1%	N/A	98.6%
Austria	71.7%	63.0%	49.0%	93.2%
Australia	99.1%	38.3%	N/A	97.6%
Czech Republic	98.9%	29.7%	75.1%	97.4%
Slovenia	84.2%	63.0%	48.2%	99.9%
Finland	70.8%	17.3%	55.9%	95.1%

Continued Table 3. *Standardization indicators among OECD-defined high-performing education systems in PISA 2022. Values are percentages of students attending schools where each indicator is reported. The table includes n = 23 education systems, with the OECD average shown only as a benchmark and excluded from all analysis; the standardized-testing analysis used n = 17 education systems after excluding N/A values. Missing values marked as N/A are as follows: Netherlands, Ireland, Japan, Australia, New Zealand, and Hong Kong (China) for "Percentage of students in schools where: Students are assessed through mandatory standardized tests at least once a year." Blue = above OECD average; gray = below OECD average; colors are descriptive only.*

Education Systems	Percentage of students in schools where:			
	Teacher Mentoring is in place	Peer-to-peer tutoring is provided	Students are assessed through mandatory standardized tests at least once a year	Internal evaluations or self-evaluations are in place
Latvia	88.4%	76.6%	98.0%	100.0%
Sweden	87.6%	20.8%	100%	97.7%
New Zealand	98.2%	71.4%	N/A	100%
Macao (China)	99.6%	91.1%	89.7%	97.7%
Chinese Taipei	87.7%	78.2%	100%	98.0%
Hong Kong (China)	89.4%	67.6%	N/A	100.0%

Table 4. *Pearson correlation and simple linear regression results for associations between socio-economic fairness (SEF) in mathematics and selected system-level indicator. OECD average values were excluded from all analysis as it is not an independent high-performing education system, which is the scope of this study.*

Variables	Sample Size (n)	Pearson Correlation Coefficient (r)	Coefficient of Determination (R^2)	Regression Model
(SEF, Ability Grouping)	23	-0.1125	0.01267	y = 86.09 − 0.05712x
(SEF, Age at selection)	21	0.4995	0.2495	y = 68.33 + 1.188x
(SEF, Mandatory Standardized Testing)	17	0.4799	0.2303	y = 77.47 + 0.09948X

Variables	P-value	95% CI for Correlation r	95% CI for Slope	Slope, Standard Error (SE)
(SEF, Ability Grouping)	0.6091	[-0.5015, 0.3142]	[-0.2860, 0.1717]	0.1100
(SEF, Age at selection)	0.02115	[0.08642, 0.7660]	[0.1986, 2.178]	0.4728
(SEF, Mandatory Standardized Testing)	0.05125	[−0.0009991, 0.7805]	[-0.0006127, 0.1996]	0.04696

To examine policies that may be associated with higher SEF in mathematics, Table 3 lists some of the major policies that may correlate with self-confidence, sense of belonging, and identity. Specifically, the association between mandatory standardized testing and SEF in mathematics was analyzed to determine whether this study corroborates previous research. The Pearson correlation coefficient of (SEF, Mandatory Standardized Testing) can be shown as:

$$r_{(SEF,\ mandatory\ standardized\ testing)} = \frac{\Sigma\left(X_3 - \underline{X}_3\right)\left(Y - \underline{Y}\right)}{\sqrt{\Sigma\left(X_3 - \underline{X}_3\right)^2 \cdot \Sigma\left(Y - \underline{Y}\right)^2}}$$

Calculating the value, , indicating a moderate positive correlation between mandatory standardized testing and SEF. Using the linear regression, the fitted regression model for $r_{(SEF,\ mandatory\ standardized\ testing)}$ is $y = 77.47 + 0.09948X$

All data visualizations including Figures 1, 2, and 3 were obtained via Python coding.

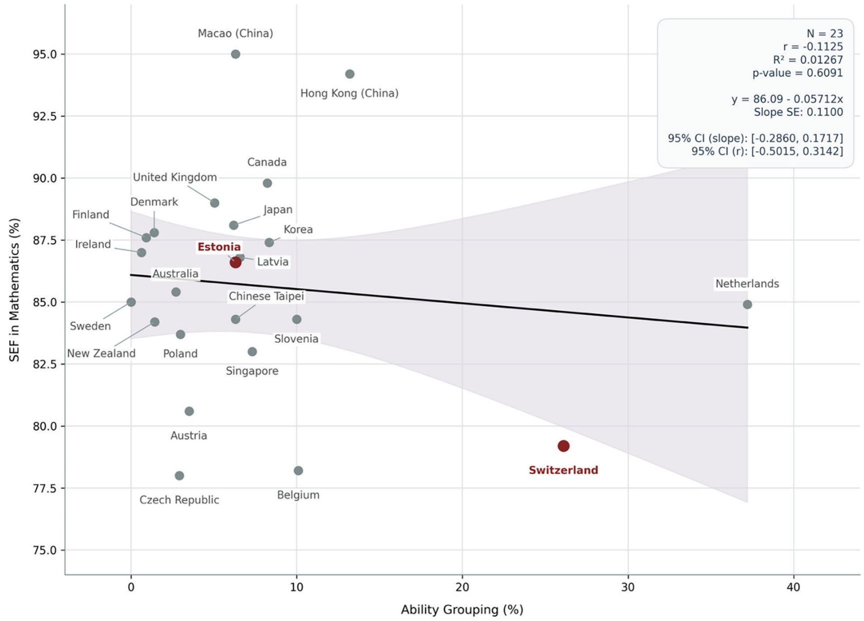

Figure 1. *Relationship between ability grouping and socio-economic fairness (SEF) in mathematics among OECD-defined high-performing PISA 2022 systems, excluding the OECD average (n = 23, r = -0.1125, R^2 = 0.01267, p = 0.6091). The fitted regression model is . The shaded region represents the 95% confidence band around the regression line. The values for y-axis "SEF in Mathematics (%)" and x-axis "Ability Grouping (%)" were derived from OECD, PISA 2022 Database, Annex B1, Chapter 4, Table B3.1.4, which correspond to the values reported in Table 1 "SEF in Mathematics" and "Percentage of students who attended schools where: students are grouped by ability into different classes for all subjects".*

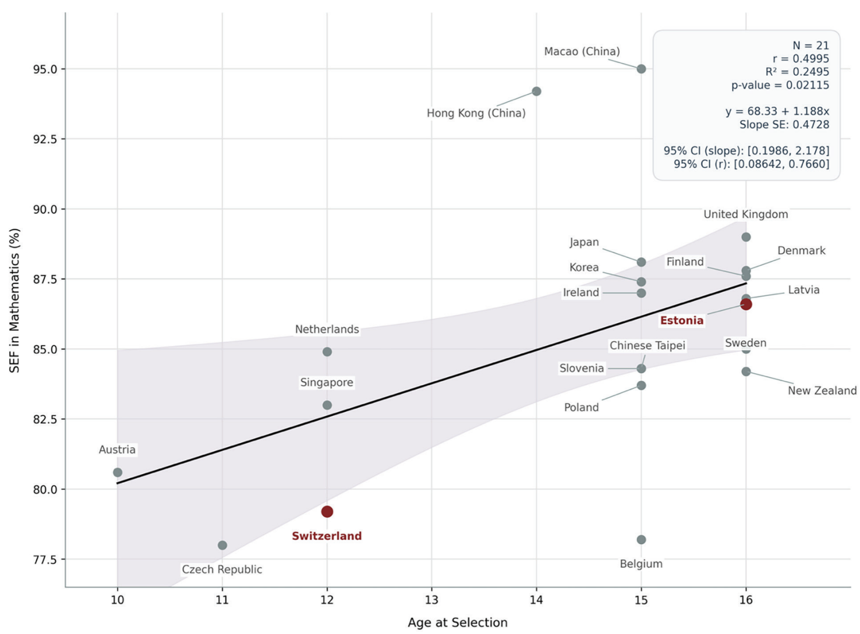

Figure 2. *Relationship between age at selection and socio-economic fairness (SEF) in mathematics among OECD-defined high-performing PISA 2022 systems, excluding the OECD average and systems with unavailable age-selection data (n = 21, r = 0.4995, R^2 = 0.2495, p = 0.02115). The regression model is The shaded region represents the 95% confidence band around the regression line. The values for y-axis "SEF in Mathematics (%)" and x-axis "Age at selection" were derived from OECD, PISA 2022 Database, Annex B1, Chapter 4, Table B3.1.4, which correspond to the values reported in Table 1 "SEF in Mathematics" and "First age at selection in the education system."*

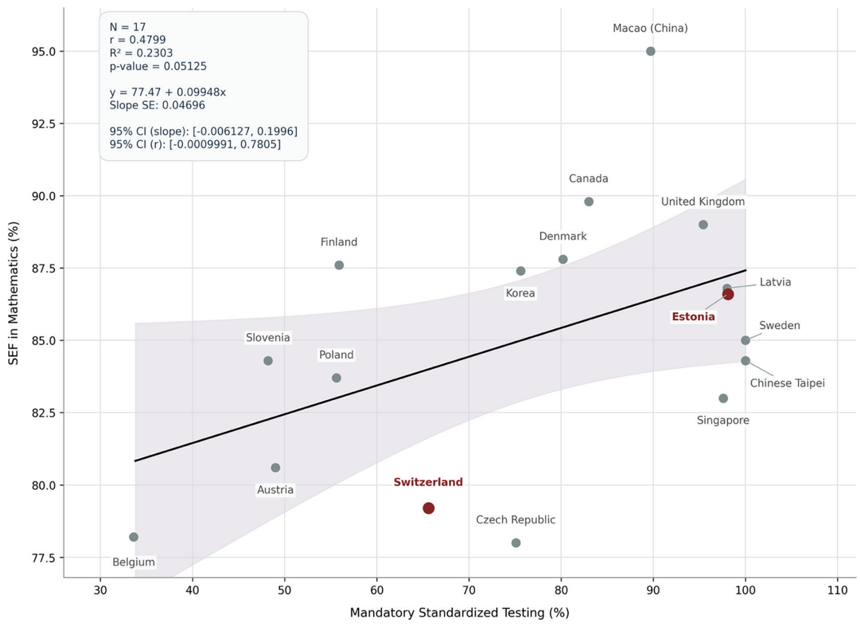

Figure 3. *Relationship between mandatory standardized testing and socio-economic fairness (SEF) in mathematics among OECD-defined high-performing PISA 2022 systems, excluding the OECD average and systems with unavailable standardized-testing data (n = 17, r = 0.4799, R² = 0.2303, p = 0.05125). The fitted regression model is . The shaded region represents the 95% confidence band for the fitted regression line. The values for y-axis "SEF in Mathematics (%)" were derived from OECD, PISA 2022 Database, Annex B1, Chapter 4, Table B3.1.4, which correspond to the values reported in Table 1 "SEF in Mathematics." The values for x-axis "Mandatory Standardized Testing (%)" were derived from OECD, PISA 2022 Database, Annex B1, Chapter 6, Table II.B1.6.39, which corresponds to the values reported in Table 3 "Students are assessed through mandatory standardized tests at least once a year."*

RESULTS AND DISCUSSION

After computing the correlation values and line of regression, the findings of the data are discussed. Firstly, this study calculated the Pearson correlation coefficient of SEF in mathematics and the percentage of ability grouping, and SEF in mathematics and age at selection. The values are -0.1125 and 0.4995, respectively. Comparing the two values, there is a stronger correlation between SEF and age at selection than between SEF and ability grouping. For the correlation between SEF and age at selection, the R^2 value of 0.2495 and a p-value of 0.02115 substantiate that the correlation is statistically significant at a 5 percent level and explains 24.95% of the variation in SEF across the systems. Furthermore, as Figure 2 shows, Estonia has a later age at selection than Switzerland, and correspondingly a higher SEF as well. Thus, the association of SEF and later age at selection supports hypothesis 1, specifically that *earlier age at selection is associated with lower SEF in mathematics.*

However, for the correlation between SEF and ability grouping, the Pearson correlation coefficient is -0.1125, indicating a weak negative correlation. In addition, the p-value greater than 0.05 at 0.6091 means that the correlation is not statistically significant, and the R^2 value at 0.01267 means that ability grouping only explains 1.267% of the variation in SEF. While Figure 1 illustrates Switzerland with higher ability grouping and a lower SEF than Estonia as hypothesized, the calculations of Figure 1 and Table 4 indicate that the correlation between SEF and ability grouping does not empirically support hypothesis 1. Indeed, Switzerland deviates from the fitted regression model of (SEF, ability grouping) by 6.209%, signaling that there may be external factors that could account for the SEF in Switzerland. The 95% confidence interval for the slope included 0, further indicating that the association between SEF and ability grouping is inconclusive in this sample. Therefore, hypothesis 1 remains only partially supported by the correlation between SEF and the age at selection.

Exploring further, the Pearson correlation coefficient between SEF and mandatory standardized testing is calculated at 0.4799, indicating a moderate positive correlation. The p-value of 0.05125 is just above the conventional 0.05 threshold, indicating that the association approached but did not reach statistical significance. The R^2 value at 0.2303 indicates that mandatory standardized testing explains 23.03% of the variation in SEF. However, the 95% confidence interval narrowly included 0. Consequently, the values calculated suggest that more prevalent mandatory standardized testing may be associated with higher SEF but should be interpreted cautiously and only treated as suggestive. In the case of Switzerland and Estonia, where mandatory standardized testing is at 65.6% and 98.1%, respectively, the SEF is higher in Estonia at 86.6% than the 79.2% of Switzerland, aligning with hypothesis 3. This aligns with the past literature by Bishop and Mane (6). Nevertheless, considering that Switzerland deviated from the fitted regression model by 4.796%, it suggests that there may be additional factors other than mandatory standardized testing in affecting SEF. Ultimately, the correlation between SEF and mandatory standardized testing partially substantiates hypothesis 3, but not conclusively: *more prevalent mandatory standardized testing is associated with higher SEF in mathematics.*

Exploratory research question 2 asked whether *higher stratification is associated with lower disadvantaged students' self-confidence and sense of belonging.* Because this question was assessed descriptively through comparison of only Switzerland and Estonia, the evidence is interpreted as exploratory rather than a tested cross-system association. Despite Estonia possessing higher accountability and school-level indicators compared to Switzerland, which include peer-to-peer tutoring, teacher mentoring, standardized testing, and internal/self-evaluation in Table 2 and 3, Switzerland exhibits a higher index of sense of belonging at 0.36 and self-confidence in education at 64.8% than those of Estonia at -0.14 and 56.3%. Therefore, the data, with a plausible combination of external factors, do not establish that higher stratification is associated with lower disadvantaged students' self-confidence and a lower sense of belonging. The isolation index for disadvantaged students was lower in Switzerland than in Estonia, with values of 0.15 and 0.17 respectively. On the other hand, the values for advantaged students were higher in Switzerland than in Estonia, at 0.20 and 0.18 respectively. Nevertheless, the differences were small and interpreted descriptively. Overall, the data do not substantially support that stratification is associated with lower self-concept. Exploratory research question 2 remains inconclusive.

Furthermore, descriptively analyzing data across Switzerland and Estonia, the change in mathematics performance across disadvantaged students and advantaged students show that Switzerland has a net difference of 17 while Estonia has a net difference of 13 (|*disadvantaged students-advantaged students*|). The change in mathematics performance for disadvantaged students in Estonia is at -15, less than the -23 of that of Switzerland. This suggests that disadvantaged students in Switzerland relatively experienced a larger decline than Estonia. While this may be due to the larger stratification in Switzerland, this trend alone is insufficient to establish that higher stratification translates directly into lower academic performance. Thus, exploratory research question 2 remains inconclusive.

Limitations

Although this research may suggest a correlation, it is not causal. In other words, external factors cannot be ignored when considering the implications. To note a few, the data were exclusively taken from OECD PISA 2022, meaning that there may have been important variables absent in the source. Furthermore, this research specifically analyzes SEF and omits key variables, such as gender inequality, the existence of external tutoring, and immigration. Thus, it is possible that the variables not taken into consideration may have played a role in determining the difference in SEF in Switzerland and Estonia, rather than solely due to ability grouping and age at selection in education.

Regarding Figure 1, 2, and 3, the sample size was only among the high-performing education systems: Switzerland, Estonia, New Zealand, Slovenia, Korea, Japan, Australia, Austria, United Kingdom, Latvia, Sweden, Finland, Czech Republic, Netherlands, Belgium, Ireland, Canada, Singapore, Denmark, Macao (China), Chinese Taipei, Hong Kong (China), and Poland (n = 23); not in particular order. The relatively small sample size limits the validity of the established trends. Moreover, certain education systems come with missing points (e.g., the Netherlands, Ireland, and Japan in "Percentage of students in schools where students are assessed through mandatory standardized tests at least once a year"; Canada in "first age at selection in the education system"). This affects the validity of the calculated regression model and may introduce bias to the data. The findings are descriptive rather than statistically definitive.

Exploratory research question 2 was assessed descriptively through the comparison between

Switzerland and Estonia, rather than through regression analysis. Therefore, the findings regarding self-confidence, sense of belonging, isolation, and change in mathematics performance are inconclusive and inapplicable to any other education system than Switzerland and Estonia.

This study utilizes published data from OECD PISA 2022. Because the data are system-level estimates rather than individual student-level microdata, the calculated values, such as the Pearson correlation coefficient, may include uncertainties that were not included in the data. Thus, the reported p-values and confidence intervals may understate the full level of uncertainty, as plausible values, replicate weights, or OECD standard errors for each estimate were not incorporated due to unavailability.

CONCLUSION

This research examines the association of age at selection and ability grouping with socio-economic fairness, as shown by Switzerland and Estonia. The findings suggest that earlier age at selection is associated with lower SEF in mathematics, while the association between ability grouping and SEF remains inconclusive. Second, more prevalent mandatory standardized testing is partially associated with higher SEF. Third, the exploratory evidence does not establish that higher stratification necessarily translates into lower self-confidence, sense of belonging, or lower academic performance in mathematics; it remains inconclusive. These findings demonstrate that education inequality is not solely a trade-off for high performance, but also how education systems are structured and organized.

From a policy perspective, governments struggling to increase SEF in mathematics may seek to not only reduce stratification within education systems but also strengthen accountability systems such as mandatory standardized testing, which introduce a uniform benchmark for all schools and students. This would mean a more uniform standard and education progress within and between schools.

Given the scope of this research, which focuses on high-performing education systems in mathematics, the associations of ability grouping, age at selection, and mandatory standardized testing with SEF in mathematics may differ for median- to low-performing education systems in mathematics. Thus, future research should focus on median or low-performing education systems, as well as the association with other stratification policies. Furthermore, gender inequality and immigrant students in education should also be included within the scope of research on socio-economic fairness, as there is sufficient data in the OECD PISA 2022. Due to the absence of data regarding how each school conducted ability grouping, there may be inconsistencies even within countries with similar percentages of students going to schools with ability grouping (e.g., an extreme extent of ability grouping and low ability grouping). Future work will continue upon how individual schools may differ in the SEF or student wellbeing with different levels of ability grouping.

ACKNOWLEDGEMENT

The author would like to thank my mentor Joshua Berry for leading and reviewing this paper.

CONFLICT OF INTEREST

The author declares that there are no conflicts of interest related to this work.

REFERENCES

1. Organization for Economic Co-operation and Development. Education equity [Internet]. Paris: OECD. Available from: https://www.oecd.org/en/topics/education-equity.html (accessed on 2026-03-15)
2. Boliver V, Schindler S. Introduction to "Mechanisms of Educational Stratification." *Res Soc Stratif Mobil*. 2021 Dec; 76: 100641. https://doi.org/10.1016/j.rssm.2021.100641
3. Dickinson EE. Coleman Report set the standard for the study of public education [Internet]. *Johns Hopkins Mag*. 2016 Winter. Available from: https://hub.jhu.edu/magazine/2016/winter/coleman-report-public-education/ (accessed on 2026-06-19)
4. Coleman JS, Johns Hopkins University. Equality of Educational Opportunity (COLEMAN) Study (EEOS), 1966 [Internet]. Ann Arbor (MI): Inter-university Consortium for Political and Social Research; 2007. Available from: https://doi.org/10.3886/icpsr06389.v3 (accessed on 2026-06-19)
5. Organization for Economic Co-operation and Development. Student socio-economic status [Internet]. Paris: OECD; 2022. Available from: https://www.oecd.org/en/topics/sub-issues/student-socio-economic-status.html (accessed on 2026-06-19)
6. Bishop JH, Mane F. Educational reform and disadvantaged students: are they better off or worse off? *SSRN Electron J*. 2004 Nov 5. https://doi.org/10.

2139/ssrn.615566
7. Schofield JW. International evidence on ability grouping with curriculum differentiation and the achievement gap in secondary schools. *Teach Coll Rec*. 2010 May; 112 (5): 1492–528. https://doi.org/10.1177/016146811011200506
8. Jetten J, Peters K. The social psychology of inequality. Cham: Springer; 2019. https://doi.org/10.1007/978-3-030-28856-3
9. Horn D. Age of selection counts: a cross-country analysis of educational institutions. *Educ Res Eval*. 2009 Jul; 15 (4): 343–66. https://doi.org/10.1080/13803610903087011
10. Haynes NM. National centralization of education: a comparative analysis of global education systems [Internet]. ResearchGate; 2025 Mar. Available from: https://doi.org/10.13140/rg.2.2.30288.83202 (accessed on 2026-06-19)
11. Blanco-Varela B, Amoedo JM, Sánchez-Carreira MC. Analysing ability grouping in secondary school: a way to improve academic performance and mitigate educational inequalities in Spain? *Int J Educ Dev*. 2024 May; 107: 103028. https://doi.org/10.1016/j.ijedudev.2024.103028
12. Organization for Economic Co-operation and Development. PISA 2022 database [Internet]. Paris: OECD; 2022. Available from: https://www.oecd.org/en/data/datasets/pisa-2022-database.html (accessed on 2026-03-15)
13. Ireson J, Hallam S, Plewis I. Ability grouping in secondary schools: effects on pupils' self-concepts. *Br J Educ Psychol*. 2001 Jun; 71 (2): 315–26. https://doi.org/10.1348/000709901158541
14. Gutiérrez N, Mercader I, Carrión JJ, Trigueros R. Self-concept and feeling of belonging as a predictor variable of the attitude towards the study from the PISA 2018 report. *Educ Sci*. 2022 Feb; 12 (2): 91. https://doi.org/10.3390/educsci12020091
15. Harrold K, Fett AK, Haenschel C. The effects of social identity incompatibility on student mental health. *Educ Sci*. 2024 Dec; 14 (12): 1407. https://doi.org/10.3390/educsci14121407
16. Fleischmann M, Hübner N, Nagengast B, Trautwein U. The dark side of detracking: mixed-ability classrooms negatively affect the academic self-concept of students with low academic achievement. *Learn Instr*. 2023 Aug; 86: 101753. https://doi.org/10.1016/j.learninstruc.2023.101753
17. Avvisati F, Wuyts C. The measurement of socio-economic status in PISA. *OECD Education Working Papers*. 2024 Sep 19. No. 321. https://doi.org/10.1787/0c5b793c-en
18. Dustmann C. Parental background, secondary school track choice, and wages. *Oxf Econ Pap*. 2004 Apr; 56 (2): 209–30. https://doi.org/10.1093/oep/gpf048
19. Xie Y. Education systems and social background inequality of educational opportunity: new evidence. *Int J Educ Res*. 2026; 138: 102975. https://doi.org/10.1016/j.ijer.2026.102975
20. Eesti Elu. EDU-cation: comparing Estonian education on an international level [Internet]. Eesti Elu; 2023 Dec 11. Available from: https://eestielu.ca/edu-cation-comparing-estonian-education-on-an-international-level/ (accessed on 2026-06-19)
21. Organization for Economic Co-operation and Development. Estonia: OECD global action on promoting social and solidarity economy ecosystems [Internet]. Paris: OECD; 2023. Available from: https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/social-economy-and-social-innovation/country-fact-sheets/country-fact-sheet-estonia.pdf (accessed on 2026-06-19)
22. Organization for Economic Co-operation and Development. Switzerland: foundations for growth and competitiveness 2026 [Internet]. Paris: OECD; 2026 Apr 9. Available from: https://www.oecd.org/en/publications/2026/04/foundations-for-growth-and-competitiveness-2026_f68a156b/full-report/switzerland_f52df388.html (accessed on 2026-07-29)