International Journal of Educational Development 107 (2024) 103028

Contents lists available at ScienceDirect

International Journal of Educational Development

journal homepage: www.elsevier.com/locate/ijedudev

Analysing ability grouping in secondary school: A way to improve academic performance and mitigate educational inequalities in Spain?

Bruno Blanco-Varela [a], José Manuel Amoedo [a], María Carmen Sánchez-Carreira [a,b,*]

[a] ICEDE Research Group, Applied Economics Department, Faculty of Economics and Business Sciences, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[b] CRETUS, Santiago de Compostela, Spain

ARTICLE INFO

JEL codes:
I21
I24
I25
I26

Keywords:
Ability grouping
Socio-educational vulnerability
Efficiency
Educational inequalities
Educational economics
Propensity score matching

ABSTRACT

Socio-educational vulnerability refers to the circumstances in which students face difficulties or challenges arising from socio-economic or contextual factors, such as a lack of resources or adverse family environments. Ability grouping during secondary school is widely used to improve student performance. The methodology tailored to students performance levels aims at providing improvements in their scores. This paper explores the complex relationship between ability grouping and socio-educational vulnerability. More specifically, it investigates how different student groups are affected by attending schools that implement ability grouping as compared to those that do not. The study is conducted in Spain, using the Programme for International Student Assessment (PISA) database and Propensity Score Matching methodology. The findings suggest that the ability grouping policy hinder equal opportunities for academically disadvantaged students, resulting in decreasing their abilities.

1. Introduction

Improving academic performance is a cornerstone in the assessment of educational systems. Educational achievement and the quality of learning are crucial indicators that reflect the effectiveness of an educational system. Likewise, effectively addressing school failure is crucial, as it determines the quality of education and the future of students. Improving academic performance, especially among low-achieving and socio-educationally vulnerable students, is key to ensure equal opportunities and increase the competitiveness of economies in the knowledge era. It is requested to implement policies that allow for the identification of areas for improvement and the development of strategies to ensure quality education and a successful future for students.

Ability grouping is considered as a strategy aimed at achieving higher efficiency for all students by placing those with similar skills in classes tailored to their abilities. However, it is noteworthy that there are opposing arguments to this practice. Literature shows that a student performance does not only depend on their initial abilities but also on the class average ability, apart from being influenced by family, economic factors, and the school demographic composition.

The main aim of this paper is to study the effects of ability grouping in 15-year-old secondary school students in terms of efficiency and equity based on socio-educational vulnerability in Spain. Specifically, two main questions are posed. Firstly, it is raised whether ability grouping has significantly positive effects on student performance, regardless of their socio-educational vulnerability. Secondly, it is analysed whether this methodology benefits students with higher socio-educational vulnerability to a higher extent, reducing the gap between good and low achievers.

One of the main novelties of this study lies in delving into the heterogeneous effects of ability grouping. In this way, it provides additional evidence on the controversy regarding the effects of this educational methodology in a specific educational context. The Spanish case is relevant due to the high rates of school failure and the need for implementing educational policies that may lower it. Furthermore, this study presents two methodological novelties. The first one refers to the formulation of a synthetic index to measure the socio-educational vulnerability from each student. The second one concerns the use of Propensity Score Matching to analyse the effects of ability grouping depending on the socio-educational vulnerability.

The used database is PISA compiled by the OECD every three years,

* Correspondence to: Faculty of Economics and Business Sciences, Santiago de Compostela (A Coruña), Avenida Burgo das Nacións s/n, 15782, Spain.
E-mail address: carmela.sanchez@usc.es (M.C. Sánchez-Carreira).

https://doi.org/10.1016/j.ijedudev.2024.103028
Received 2 November 2023; Received in revised form 8 February 2024; Accepted 19 March 2024
Available online 4 April 2024

and the 2018 year is under study because it is the most recently published one. Then, the Propensity Score Matching (PSM) method is employed to establish treatment and control groups that are balanced in relevant characteristics. PSM is used to ensure that the groups are comparable and that differences in outcomes can be attributed to the relevant variable, which is socio-educational vulnerability in this case.

This paper is organised into four sections apart from this introduction. The second section presents the literature review on the relationship between ability grouping and socio-educational vulnerability. The third section describes the database and the methodology. The fourth section deals with the results of the analysis, focusing on the effects of ability grouping in Spain. The final section shows the main conclusions and implications of the research.

2. Socio-educational vulnerability in the context of ability grouping: a close link

Ability grouping is considered as a strategic measure implemented to boost overall student performance improvement in educational systems. It is aligned with the identification of differences in learning and it stresses the improvement of low-performing or highly educationally vulnerable students. In this way, the distribution of different performance levels provides each student with the opportunity to reach their maximum potential.

This educational methodology can be applied between classes or within the same classes, as well as within the same school or between schools (Becker et al., 2014). When the grouping is between classes, the group of students in a course is divided into segments based on the performance level: resulting in a reference group and a group of low-performing students. The purpose of this division is to enable the low-performing group to improve and reach the same level of competences as the reference group (Aramendi et al., 2012; Pekkarinen, 2018). Both groups have curriculum homogeneity but use different teaching practices. This enhanced adaptation could potentially achieve improvements in the quality of education, extending beyond basic skills in reading or mathematics, and leading to a better average academic performance. The grouping of skills between schools is not common in the Spanish context, due to the general geographical criteria for enrolment, prioritizing the proximity of the student to the school (Bonal and Bellei, 2018).

The decision regarding ability grouping is made by the school on a discretionary basis. The composition of the groups is typically determined by the teaching team or school administration in most cases. This methodology applied within the same classroom refers to the frequency with which teachers have students learn with peers of similar abilities. For example, high-achieving students work together at the same desk and lower-achieving students at another. When grouping is done between different classrooms, students are instructed through different methodologies (OECD, 2020). However, a third type of grouping can be established, which is the allocation of students to different schools (inter-school grouping), such as Tan and Dimmonck (2020) highlight.

Tan and Dimmonck (2020) underline that ability grouping between classes may be less effective than other forms of grouping. They consider that it is not the action of grouping between classes itself that leads to the short observed effects on performance, but rather its interaction with the curriculum, classroom, and student characteristics. Therefore, the composition influences self-determination and interest in studies (Campbell, 2021; Guill et al., 2017; Caballé, 2016). The origin of self-concept is linked to processes of "labelling students." Thus, according to Campbell (2021), a discrepancy between actual ability level and revealed performance arises because ability groups are not directly based on the actual level possessed but also on revealed ability and stratified according to factors that include demographic characteristics, social hierarchies, and integrated norms. Taylor et al. (2017) consider that ability grouping misleads current educational achievement with a notion of innate academic potential. As a result of this, there may be mismatches in skill level placement that influence learners experiences, expectations, and progress.

In turn, grouping may have effect beyond performance related to the socioeconomic composition of the classroom. In this sense, numerous studies emphasize the significance of social differences and their impact on education, such as Busemeyer & Nikolai (2010); Alegre & Subirats (2013); Dupriez et al. (2012); Benito & González-Balletbò (2012); Dupriez et al. (2008); Pekkarinen (2018); Gracia & Esping-Andersen (2015); Rogero-García & Andrés-Candelas (2014). In this regard, ability grouping can negatively affect the equity of educational systems.

Teacher perceptions interfere with the assignment of student achievement levels. Therefore, teachers social biases may have repercussions in the formation of students self-concept (Francis et al., 2015; Taylor et al., 2017). Student grouping can initiate processes like a "self-fulfilling prophecy" (Francis et al., 2020), with cumulative effects. These consequences arise from the dynamic interactions of students within their environment and the perceptions and behaviours of those around them, including their teachers. In this context, Cavaco et al., 2021 underscores the role of social dynamics, such as gender, in shaping teachers expectations of their students.

Various perspectives on grouping agree that teachers expectations are significant factors influencing students learning and motivation. For instance, Uysal and Banoglum (2018) underscore the advantages of separating students to facilitate differentiated and tailored teaching approaches, especially for more homogeneous groups. On the contrary, Francis et al. (2020) and Campbell (2021) consider that performance disparities mainly stem from teachers expectations. These expectations are rooted in value judgments concerning a perceived level of demonstrated abilities, which may not always align with objective assessments.

In addition, it is crucial to underscore the role that close family and social socioeconomic circumstances play in student performance level, often associated with their socio-educational vulnerability (Aramendi et al., 2012; Benito and González-Balletbò, 2012; Fernández et al., 2017; Blanco-Varela, 2022). In this sense, socio-educational vulnerability refers to specific educational challenges due to unfavourable factors, such as low income, challenging family backgrounds, lack of access to quality educational resources, and disabilities. It can limit access to quality education and learning opportunities. Therefore, the stratification or ranking of students can be based on factors that go beyond purely academic, including socioeconomic considerations. Furthermore, if there are school segregation by achievement that provide differential education based on the socioeconomic background of students, it can lead to disparities in their educational opportunities (Aisenson et al., 2011).

Francis et al. (2015) argue that skill-based grouping cannot legitimize social stratification as a natural phenomenon. The consequences of this approach often result in higher social segregation, hindering social mobility. Students needs are not mainly driven by the requirement for tailored learning methodologies, but they are substantially influenced by socio-educational vulnerability stemming from their socioeconomic context. An unfavourable socioeconomic environment and a less inspiring academic context can interact with educators holding less ambitious expectations. In this context, mixed-ability classes can play a crucial role in supporting and attending to the most vulnerable students (Meier and Schütz, 2007; Álvarez et al., 2004; Zhang et al., 2014).

3. Methodology

The literature highlights the significant effect of individual circumstances on educational results, emphasizing the role of ability grouping as a critical factor within classroom composition. It should be noted that personal and individual background circumstances play a crucial factor in configuring educational achievement. Moreover, recent studies suggest that the practice of grouping students according to their achievement levels may generate differential effects on academic achievement, mainly considering the different levels of socio-educational vulnerability. In this sense, higher educational vulnerability can reinforce the undermining aspects of ability-grouping policies (Catalayud, 2018; Narciso, 2022). This underscores the need to carefully consider the methodological approach when investigating the complex interaction between ability grouping and socio-educational vulnerability.

The PISA report assesses the academic performance of 15-year-old students from several countries in reading, mathematics, and science, providing information on the quality of education systems. The effects of ability grouping in different classes on academic performance are analysed using PISA data from the 2018 edition for Spain, analysing schools and students[1] (OECD The Organization for Economic Cooperation and Development, 2019). The case of Spain is worth studying for two main reasons closely related to socio-educational vulnerability, which are high school dropout in comparison with other European or OECD countries, and its high level of socioeconomic inequality. It provides an opportunity to delve into an educational system featured by a substantial dropout rate and simplifies the categorization of students into broader groups based on their socio-educational vulnerability. The selection of only one country avoids the diversity among the highly diverse education systems, as found within the OECD.

PISA collects two variables about the use of ability grouping in schools. One of them reports whether the school applies it in all subjects and the other one whether it does it only for some subjects. From these two variables, the treatment variable (ABGROUP) is constructed. It gathers information regarding if a student is enrolled in a school that implements ability grouping, regardless of whether this practice applies to some or all subjects (providing values of 1 and 0, respectively).

From the data collected in PISA, the control and results variables for the analysis are built. These variables can be grouped in five dimensions, which are the following ones: individual characteristics, socioeconomic environment, individual access to learning resources, school characteristics, and academic performance. The selection of these variables is based on the treatment of achievement in the literature, which focuses on individual factors; family factors and household possessions; and ecological factors (Sicilia and Simancas,2018; Gabarró, 2010; Garrido-Yserte et al., 2019; Rogero-García and Andrés-Candelas, 2014; Agasisti and Longobardi, 2014; INEE, 2016; Owens, 2017; OECD, 2020; Martínez-Garrido and Lange, 2019).

The variables concerning individual characteristics encompass the key features of the students that may influence their academic performance. The socioeconomic background variables include the factors within a student material and family environment that may impact their academic performance. The individual access to learning resources variables describes the degree to which a student can access to individual educational resources that support and enhance their learning. The variables regarding school features assess their quality and resources that are pertinent for explaining student academic attainment. Finally, the academic performance variables comprise the overall student performance as well as their achievement in three specific knowledge areas: mathematics, science, and reading. Table 1 provides a summary of the selected variables.

The literature on the economics of education has frequently used the Socioeconomic and Cultural Index (ESCS) to study the impact of background on academic achievement, as mentioned in Section 2. This index also provides a relevant dimension of socio-educational vulnerability. However, this research has opted for a more ambitious synthetic index, which considers more dimensions of the educational ecosystem and includes the school dimension (educational resources and educational career).

Therefore, a new synthetic index is proposed to assess the socio-educational vulnerability of each student.[2] Socio-educational vulnerability has not been used very frequently in the literature, which mainly focuses on certain (individual) factors or on socio-economic vulnerability, excluding aspects linked to the student educational environment (for example, the quality of the school resources or the fact of being a repeater or not) that significantly condition student performance. The reasons for constructing this index stem from the lack of synthetic indicators in the existing literature, primarily focused on socioeconomic vulnerability, and the need to consider the available data (mainly PISA). The proposal to use a vulnerability index is grounded in the concept that vulnerability is indicative of an individual decreased ability to handle unexpected situations or contingences. However, as Blanco-Varela (2022) emphasizes, being vulnerable does not necessarily mean that such emergencies will occur. It simply suggests a lesser readiness or capacity to face them if they arise.

Concerning the calculation of the SVI, it is conducted successively as usual: first, a series of variables are weighted and aggregated to construct the sub-indicators relating to a given dimension, and then these sub-indicators are aggregated to construct the synthetic index. Some of the variables that composed the index were transformed. This happens in the case of qualitative variables and, in particular, categorical ones, such as the variables concerning the educational level of the parents, the student origin, the stage at which the students starts their studies, and the quality of the school resources (EDUFATHER, EDUMOTHER, INMIG, EDUBEGIN and RESOURCES, respectively) in order to create new variables that collects the relationship between the considered variable and vulnerability, that in some of these cases is inverse and with normalized values, as Tsaur et al. (2006) and Hermans et al. (2007) suggest. Thus, for example, higher levels of parental education imply lower vulnerability, being a native involves lower vulnerability than being a second-generation immigrant, and this, in turn, than being a first-generation immigrant; starting education at an earlier age entails lower vulnerability; and the higher the quality of educational resources presumes the lower the vulnerability. In addition, the ESCS was normalized to range between 0 and 1 and avoid negative values. A higher value in this index is associated with lower vulnerability. In the case of the modified categorical variables, the weights are assigned according to the relationship indicated by the existing literature. Thus, a higher value represents a higher vulnerability of the student in that dimension.

Table 2 describes these modifications and the origin of the new variables. In these cases, a value scale is assigned based on the specific characteristic measured. These modifications arise from the necessity to consolidate certain variables within the same dimension. This is done in accordance with the literature, considering that socio-educational vulnerability is perceived as a collection of barriers that may hinder a student access to education or the education achievements. In other words, these barriers may not inherently lead to a decline in academic performance, but they can represent a handicap.

When constructing the synthetic index, the methodology used is known as simple aggregation, giving each of the 12 considered dimensions the same weight and using variables with values between 0 and 1. The simple aggregation is the most commonly used method for building a synthetic index, being its main advantages the easy

[1] Students with missing values on any of the analysed variables are eliminated from the sample.

[2] Several OLS regressions using SVI versus the ESCS were performed to analyse their correlation with student academic performance. The results show how the SVI improves the ESCS as an explanatory factor of student academic performance both in terms of overall performance and in each of the three competencies included in PISA (Science, Reading and Mathematics).

Table 1
Summary of variables by dimension.

Dimension	Variable	Code	Description	Values
Individual characteristics	Age	AGE	Student age	15,08–16,33
	Sex	SEX	Student sex	Female
				Male
	Repeating student	REPEAT	Repeating student	Repeating
				Non Repeating
	Age at which formal education commenced	PREPRIMARY	Stage at which the students start their studies	ISCED 01
		CHILDHOOD		ISCED 02
		PRIMARY		ISCED 1
	Origin	INMIG1	Student origin	First-generation inmigrant
		INMIG2		Second-generation inmigrant
		NATIVE		Native
Socioeconomic environment	Father studies	EDUFATHER1	Level of education attained by the student father	ISCED[a] 1 & 2
		EDUFATHER2		ISCED 3 & 4
		EDUFATHER3		ISCED 5
		EDUFATHER4		ISCED 6, 7, 8 & 9
	Mother studies	EDUMOTHER1	Level of education attained by the student mother	ISCED 1 & 2
		EDUMOTHER2		ISCED 3 & 4
		EDUMOTHER3		ISCED 5
		EDUMOTHER4		ISCED 6, 7, 8 & 9
	Social, economic and cultural status	ESCS	Social, economic and cultural status index (ESCS) from the student	-3195–3.611
	Own desk	OWNDESK	Own desk in home	Own desk
	Own bedroom	OWNBEDROOM	Own bedroom in home	Own bedroom
	Own space	OWNSPACE	Own quiet study space in home	Own study space
Individual acces to learning resources	Computer	PC	Disposable computer at home to study	Computer
	Internet access	INTERNET	Internet acces at home	Internet
	Books at home	BOOKS1	Number of books in the student home	< 10 books
		BOOKS2		11–25 books
		BOOKS3		26–100 books
		BOOKS4		101–200 books
		BOOKS5		201–500 books
		BOOKS6		> 500 books
School characteristics	Discipline	DISCLIMA	Disciplinary climate in test language lessons in the school	-2.712–2.035
	Region	REGION	City or autonomous community to which the school belongs	Spanish NUTS 2
	School ownership	CHARTER	School ownership	Charter
		PUBLIC		Public
		PRIVATE		Private
	Municipality size	MSIZE1	Number of inhabitants of the municipality in which the school is located	< 3000 inhabitants
		MSIZE2		3001–15,000 inhabitants
		MSIZE3		15,001–100,000 inhabitants
		MSIZE4		100,000–1000,000 inhabitants
		MSIZE5		> 1000,000 inhabitants
	Class size	CLSIZE1	Average school class size	< 15 students
		CLSIZE2		16–20 students
		CLSIZE3		21–25 students
		CLSIZE4		36–30 students
		CLSIZE5		31–35 students
		CLSIZE6		36–40 students
		CLSIZE7		41–45 students
		CLSIZE8		46–50 students
		CLSIZE9		> 50 students
	Students by teacher ratio	STRATIO	Students by teacher ratio	1–51,579
	School resources	SCHRESOURCES1	School inadequate or poor quality educational material	A lot
		SCHRESOURCES2		To some extent
		SCHRESOURCES3		Very little
		SCHRESOURCES4		Not at all
	Assisting staff lack	SCHSTAFF1	Lack of assisting staff	A lot
		SCHSTAFF2		To some extent
		SCHSTAFF3		Very little
		SCHSTAFF4		Not at al
	Computers by student ratio	PCRATIO	Computers by student ratio	0–25
Academic performance	Global performance	GLOBALPERF	Average of plausible values in Global Competence	218.7–801.6
	Science performance	SCIENCEPERF	Average of plausible values in Science	193.7–767.9
	Reading performance	READINGPERF	Average of plausible values in Reading	176.4–740.6
	Mathematics performance	MATHPERF	Average of plausible values in Mathematics	185.4–733.0

Source: Own elaboration based on OECD The Organization for Economic Cooperation and Development, 2019

[a] It refers to the levels of the International Standard Classification of Education (ISCED) to classify educational levels in different countries. It ranges from ISCED 1, corresponding to primary education to ISCED 9 for the advanced tertiary education, including doctoral programs or other higher education levels.

Table 2
Modification of variables used to calculate the SVI.

Variable Code	Description
ESCSN	ESCS variable standardised between 0 and 1
EDUFATHER	= EDUFATHER1*1 + EDUFATHER2*0.66 + EDUFATHER3*0.33 + EDUFATHER4*0
EDUMOTHER	= EDUMOTHER1*1 + EDUMOTHER2*0.66 + EDUMOTHER3*0.33 + EDUMOTHER4*0
INMIG	= NATIVE*0 + INMIG1*0.5 + INMIG*1
EDUBEGIN	= PRIMARY*1 + CHILDHOOD*0,5 + PREPRIMARY*0
RESOURCES	= SCHRESOURCES1*1 + SHCRESOURCES2*0.66 + SCHRESOURCES3*0.33 + SCHRESOURCES4*0

Source: Own elaboration

interpretation and calculation.[3] Thus, the SVI is calculated based on the following expression:

$$SVI = \left[1 - \left(\frac{ESCSN + OWNDESK + OWNBEDROOM + OWNSPACE + PC + INTERNET}{6}\right)\right]^{1/2} + \left[\frac{RESOURCES + EDUFATHER + EDUMOTHER + INMIG + EDUBEGIN + REPEAT}{6}\right]^{1/2}$$

The SVI values range between 0 (the minimum vulnerability) and 1 (maximum vulnerability). Fig. 1 represents the distribution of students in the sample according to their SVI value density function. Three groups are formed to divide the students in the sample into groups according to their socio-educational vulnerability. The first group consists of low vulnerability students (SVI below the 25th percentile value), the second one groups medium vulnerability students (SVI between the 25th and 75th percentile, both included) and the third group embraces high vulnerability students (SVI above the 75th percentile). Besides allowing to categorize students close to the mode as those with medium vulnerability, this division facilitates the creation of sufficiently large groups to get good matches. The two vertical asymptotes in Fig. 1 represent these divisions: the first asymptote shows the 25th percentile (SVI = 0.141) and the second one the 75th percentile (SVI = 0.276). Thus, the sample is divided into three groups, which will be separately analysed below. It should be noted that, when examining the case of Spain, the group labels become relevant within that particular context. In other words, while students with high vulnerability in Spain may not show the same level of vulnerability in less developed countries, their vulnerability is relative within the Spanish context. Although Spain, as a developed country, shows an overall low socio-educational vulnerability in comparison with other countries, this does not prevent the existence of high rates of school failure and strong internal inequalities. It should also be noted that the exposure to high vulnerability has subsequent consequences, as labour market insertion or the own quality of employment show (Blanco-Varela, 2022).

In order to assess the impact of ability grouping, it is crucial to compare the performance of individuals who have experienced ability grouping with those who have not. Nevertheless, given that these groups may show disparities in other factors that can influence academic performance, it becomes imperative to employ a methodology that can mitigate these variations.

Propensity Score Matching (PSM) is used to obtain balanced treatment and control groups (Rosenbaum and Rubin, 1983). Specifically, different combinations of the nearest neighbour methodology (NNM) and exact matching are applied (Guill et al., 2017; Stuart, 2010; Zhang et al., 2014), which allows for more accurate matches. The combination of both methodologies enables the selection of those variables where exact matching of individuals is most important (Stuart et al., 2011). In the present case, these variables are REPEAT, SEX, ORIGIN, and REGION. In the last case, the inclusion of the variable REGION is due to the fact that education competences rest on the regional level. Thus, it may be differences among each reality. The other variables belong to the group of student basic characteristics. Apart from the variables used for exact matching abovementioned, the other variables in Table 1 are considered as non-exact matching. Moreover, the SVI and its components are included in the matching. In this way, matchings are obtained with similar SVI and, in turn, as similar as possible or identical, in their components. This is only obtained by including both together.

In order to obtain the best possible balancing to eliminate endogeneity problems, different ratios (1, 3, 5 and 10), distance "glm", replacement and discard are applied.[4] Standardised bias and pseudo-R^2, and graphical analysis are used to study and to test the balance. Table 3 shows the overall balance measures for each group of individuals (high, medium, and low vulnerability) before and after each type of matching. In addition, Tables A.1, A.2 and A.3 in the Appendix collects the variables individual standardised bias for each vulnerability group.

Maintaining a strong balance between the treatment and control groups facilitates the assessment of the impact of ability grouping by directly calculating the difference between these two groups. In this way, the weighted means for the three performance variables (overall, mathematics, science and reading) is calculated, as well as the percentage difference between the two groups as the effect of the policy. In addition, a t-test is performed to check the statistical significance of these differences.

4. Results

To measure the ability grouping effects, the Fig. 2 shows the performance percentage difference between the treatment and control group for each knowledge area and socio-economic vulnerability. Furthermore, it depicts the statistical significance of each difference. A more detailed overview of these results is given in Table A.4 in the Appendix.

[3] It should be noted that there is not an optimal method to build a synthetic index, because each alternative, mainly concerning normalization techniques and how to weight sub-indicators, presents advantages and disadvantages (den Butter and von den Eyden, 1998; Kang, 2002; Ebert and Welsch, 2004; OECD/European Union/EC-JRC, 2008). It involves that the researcher has to take decisions that entail subjectivity (Munda, 2005; OECD/European Union/EC-JRC, 2008; Saisana and Tarantola, 2002). Nevertheless, in order to check whether the simple aggregation and the weights assigned to the variables for the calculation of the synthetic index are adequate, an alternative synthetic index using Principal Component Analysis from the untransformed variables have been constructed. This latter method presents two clear disadvantages compared to the index finally used: its interpretative and calculation complexity and a lower explanatory capacity for academic performance. The tested alternatives for the index do not improve the results.

[4] In addition, it is used a 0.15 calibrator with ratio 10, which does not improve the results.

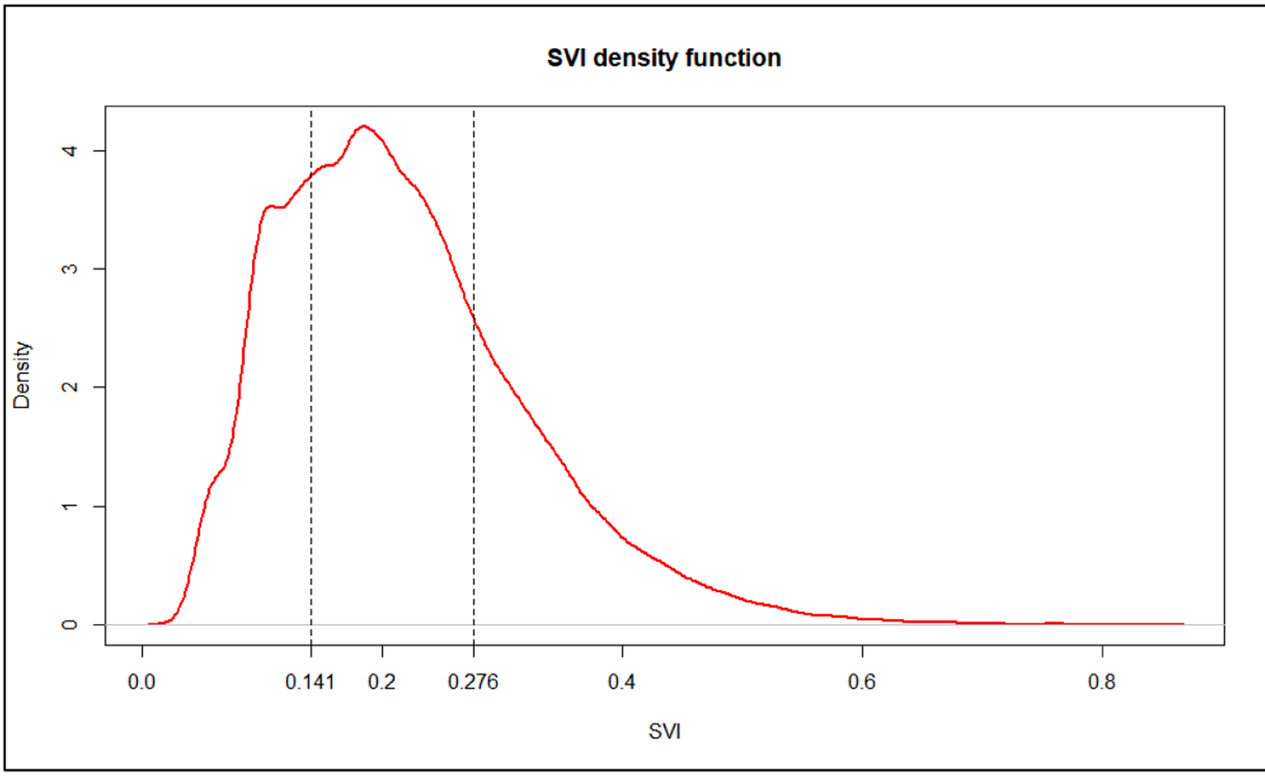

Fig. 1. SVI density function. Source: Own elaboration.

Table 3
Balance measures.

HIGH				
	Standardised bias	Pseudo-R^2	Control individuals	Treatment individuals
Before matching	7,49%	0,0657	3.907	2.042
NN 1:1	2,82%	0,0188	2.018	1.163
NN 1:3	2,56%	0,0160	2.246	2.018
NN 1:5*	2,54%	0,0153	2.692	2.018
NN 1:10	2,87%	0,0108	3.143	2.018
NN 1:10 C:0.15	2,36%	0,0136	2.565	1.642
MEDIUM				
	Standardised bias	Pseudo-R^2	Control individuals	Treatment individuals
Before matching	5,02%	0,0436	7.773	4.122
NN 1:1	2,92%	0,0164	2.197	4.114
NN 1:3	2,14%	0,0120	4.240	4.114
NN 1:5	1,83%	0,0102	5.324	4.114
NN 1:10*	1,67%	0,0078	6.292	4.114
NN 1:10 C:0.15	2,94%	0,0106	5.396	3.662
LOW				
	Standardised bias	Pseudo-R^2	Control individuals	Treatment individuals
Before matching	4,96%	0,0395	3.860	2.089
NN 1:1	2,82%	0,0165	1.197	2.071
NN 1:3	2,57%	0,0128	2.243	2.071
NN 1:5*	2,47%	0,0115	2.718	2.071
NN 1:10	2,66%	0,0131	3.181	2.071
NN 1:10 C:0.15	2,44%	0,0166	2.553	1.698

Note: * indicates which option is selected as the matching for the calculation of the results for each vulnerability group.
Source: Own elaboration

These results suggest that ability grouping shows predominantly negative consequences on student achievement. One particularly notable result is that vulnerable students present a significant overall decrease in performance in all three subjects, making them the group most affected by ability grouping. This is consistent with the results of Meier and Schütz (2007).

Students with medium vulnerability do not seem to benefit or be affected by ability grouping, since none of the observed effects reach statistical significance for this group. However, in the case of low vulnerability students a notable decline is observed in the area of Reading, suggesting that ability grouping may negatively affect their specific reading achievement.

Concerning the different effect by knowledge area, several aspects highlight. Reading is the competence most affected by this educational policy. Thus, negative effects are found on the performance of high and low socio-educational vulnerability students. Mathematics and Science fields share similarities, given that performance in both areas falls for individuals with high vulnerability and smoothly falls or rises for the other two groups but these findings are not statistically significant. In terms of overall achievement, the results show a drop in performance for all groups of students. However, only the decline in the high vulnerability group is statistically significant.

Therefore, the findings do not generally suggest that ability grouping presents an overall positive effect on academic performance. In this regard, socioeconomic variables appear to be more relevant than the inherent differences in learning. The socioeconomic mix of the classroom and peer effect can boost students with higher difficulties concerning performance (Ficano, 2012; Dupriez et al., 2008; Dumay & Dupriez, 2008). Furthermore, it impacts on another aspects, such as educational aspirations (Barberá et al., 2008).

The different effects observed among students with diverse levels of vulnerability highlight the need to adapt educational policies to address the tailored needs of students and attempt to achieve equitable learning outcomes. However, it seems that simply adapting the educational methodology of learning in different classrooms may not be an effective

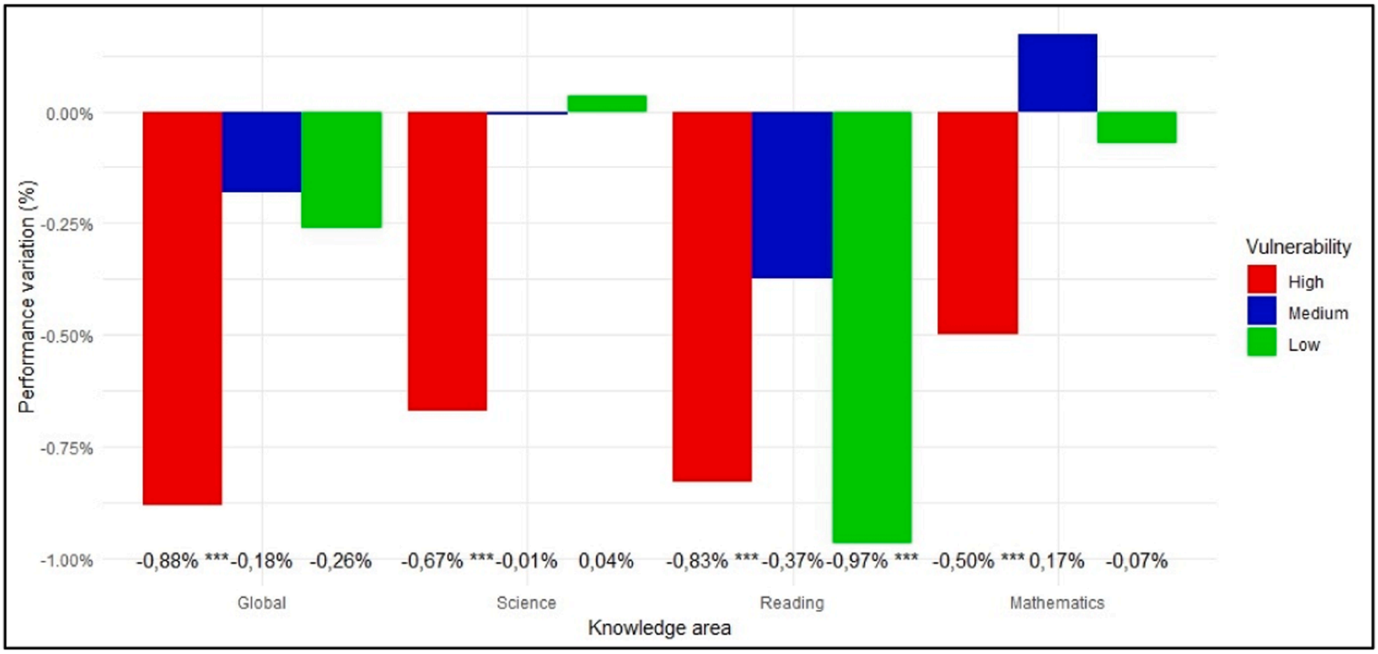

Fig. 2. Performance variation by knowledge area and socio-economic vulnerability. Note: *** p<0.01/ ** p<0.05/ * p<0.10. Source: Own elaboration based on OECD The Organization for Economic Cooperation and Development, 2019.

means of improving achievement. This highlights the need for alternative measures that focus on diversifying the composition of classrooms and providing comprehensive support to highly vulnerable families.

5. Conclusions

Ability grouping is a strategic methodology used in educational systems to improve overall student performance by grouping them according to their skills and needs. Through an adapted methodology, the low-performing students were able to achieve the same competencies as the reference group. However, ability grouping can worsen inequalities, especially when group composition is reinforced by socioeconomic factors, leading to social stratification, and limiting social mobility. In this context, the effect of socio-educational vulnerability on limiting educational outcomes could be widened by the composition of the group, creating a reinforcement gap. Also note that grouping processes are affected by social biases and prejudices. Thus, their application depends on teachers' expectations about students' learning potential, creating self-fulfilling prophecies.

This paper analyses the effects of ability grouping on the performance of students in Spanish secondary schools based in PISA-2018. Specifically, it studies the effects on the achievement of different groups of students according to their socio-educational vulnerability, examining whether ability grouping can have differential effects on low, medium, and high vulnerability students. To classify students according to their socio-educational vulnerability, a synthetic index, is calculated. Propensity Score Matching is used to obtain similar treatment and control groups, isolating the effect of ability grouping. This enables to estimate the effect of ability grouping as the difference between the treatment and control groups.

The outcomes of ability grouping in the study show consistently adverse and regressive effects. More precisely, the findings reveal substantial negative impacts across diverse knowledge domains, affecting students with varying levels of vulnerability. These detrimental effects are not counterbalanced by any corresponding positive outcomes. Moreover, it should be noted that high level vulnerability students are those who were the most affected by these negative consequences. In fact, it is the only group that shows a general decline in academic performance. Considering these findings, the adoption of ability grouping as a teaching methodology seems ineffective.

The findings suggest that classroom diversity and peer interactions can have a positive influence on the performance of students facing learning challenges, particularly when such policies offset social barriers limiting the educational opportunities of those with fewer advantages. Thus, mixing classrooms with students of different achievement levels is useful to foster the academic growth of lower performers. Heterogeneous or mixed-ability grouping has several positive effects on the efficiency and equity of education systems. First, it facilitates an environment in which students of diverse abilities can interact and collaborate. This interaction allows high-achieving students to serve as mentors or role models to their peers, helping them if they face academic difficulties. Second, mixed-ability classrooms can help reduce the stigma associated with low achievers. When students of different abilities are mixed, it is less obvious who differs in difficulty than when students are segregated in another classroom.

In this context, mixed teaching stands out as an effective approach, as supported by research on the composition effect. Numerous studies, such as Ficano (2012) and Dupriez et al. (2008), (2012), underline the relevance of diversity in the classroom. These studies highlight that the classroom environment does not only facilitate knowledge sharing, but also reinforces aspirations and strengthens social cohesion.

Three aspects concerning the PISA data should be noted as main limitations of this study. The first issue is that it does not indicate the knowledge area in which ability grouping is applied. Thus, it simply states whether it is applied in all subjects or only in some of them. The second matter refers to the lack of enough information whether the grouping takes place within or between classrooms. The third aspect concerns the focus of this paper on ability grouping within school, neglecting inter-school grouping. Despite knowing the relevance of inter-school grouping, this weakness is not significant for this current study in Spain due to two main reasons: 1) the availability of PISA database, which does not allow a direct analysis based on performance, whereas it might enable an analysis based on the enrolment of students from different socio-economic backgrounds in specific schools; and 2) inter-school practice is not allowed in Spain (at least in the public system), as student enrolment is generally based on geographical criteria (proximity to the school) and not on academic record. Future extensions to this study range from a comparative analysis of the OECD countries to case studies comparing particular schools.

Authors statements

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

They repect the ethics in publishing and follow the COPE guidelines.

CRediT authorship contribution statement

José Manuel Amoedo: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **María Carmen Sánchez-Carreira:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Bruno Blanco-Varela:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Appendix A

Table A1
High socio-educational vulnerability bias.

Variable	Balancing measures: standarised bias					
	Before matching	NN1:1	NN1:3	NN1:5	NN1:10	NN1:10 C:0.15
SVI	17,67%	-0,45%	0,82%	1,12%	3,13%	-2,33%
AGE	4,55%	-1,60%	0,19%	-0,47%	-0,79%	-5,59%
SEX	-4,97%	0,00%	0,00%	0,00%	0,00%	0,00%
REPEAT	6,16%	0,00%	0,00%	0,00%	0,00%	0,00%
PREPRIMARY	2,32%	-7,56%	-4,03%	-2,86%	-2,77%	-5,55%
CHILDHOOD	-3,36%	7,45%	3,33%	2,26%	1,83%	5,42%
PRIMARY	3,68%	0,79%	2,75%	2,34%	3,56%	0,77%
IMMIG1	5,75%	0,00%	0,00%	0,00%	0,00%	0,00%
IMMIG2	4,67%	0,00%	0,00%	0,00%	0,00%	0,00%
NATIVE	-7,63%	0,00%	0,00%	0,00%	0,00%	0,00%
EDUFATHER1	5,89%	-0,87%	0,99%	0,25%	1,75%	-2,81%
EDUFATHER2	2,33%	2,36%	0,76%	1,30%	0,68%	-0,23%
EDUFATHER3	0,74%	0,70%	1,32%	0,34%	0,72%	3,68%
EDUFATHER4	-8,25%	-1,60%	-2,77%	-1,56%	-2,91%	-0,42%
EDUMOTHER1	7,79%	0,46%	0,73%	1,21%	2,34%	-1,22%
EDUMOTHER2	5,11%	-2,94%	-1,76%	-0,91%	0,11%	-1,51%
EDUMOTHER3	-1,56%	1,38%	1,80%	0,54%	-0,01%	2,97%
EDUMOTHER4	-9,40%	0,71%	-0,76%	-0,80%	-2,12%	-0,18%
DISCLIMA	-3,50%	-2,61%	-2,08%	-3,07%	-1,67%	-0,35%
ESCS	-12,56%	-1,40%	-2,90%	-3,78%	-5,15%	0,37%
OWNDESKTOP	-1,21%	9,65%	8,46%	7,54%	4,33%	5,64%
OWNBEDROOM	-6,68%	-4,54%	-1,36%	-1,74%	-2,83%	1,18%
OWNESPACE	-4,09%	1,20%	0,07%	-0,75%	-0,78%	0,45%
PC	-2,64%	1,38%	-1,05%	-1,95%	-1,35%	1,19%
INTERNET	-0,07%	-2,84%	-0,68%	-1,43%	-1,84%	-2,63%
BOOKS1	7,22%	-6,91%	-2,68%	0,50%	0,18%	-3,55%
BOOKS2	3,11%	3,53%	1,65%	1,63%	2,45%	1,27%
BOOKS3	0,60%	1,49%	3,18%	0,42%	0,66%	2,11%
BOOKS4	-5,26%	-2,31%	-5,00%	-4,30%	-6,44%	-2,74%
BOOKS5	-5,26%	-2,31%	-5,00%	-4,30%	-6,44%	-2,74%
BOOKS6	1,75%	2,31%	0,06%	0,98%	2,72%	1,68%
REGION	-15,44%	0,00%	0,00%	0,00%	0,00%	0,00%
CHARTER	-28,27%	0,49%	-2,25%	-3,81%	-5,15%	0,70%
PRIVATE	1,58%	2,29%	-1,08%	-3,05%	-2,55%	0,90%
PUBLIC	24,66%	-1,76%	2,66%	5,21%	6,13%	-1,15%
MSIZE1	-5,55%	4,97%	2,90%	2,75%	1,56%	3,66%
MSIZE2	9,46%	1,42%	4,95%	4,68%	5,61%	2,25%
MSIZE3	-7,93%	-6,06%	-4,87%	-4,84%	-6,02%	1,54%
MSIZE4	3,94%	0,54%	-1,58%	-0,57%	-0,09%	-6,42%
MSIZE5	-6,13%	4,55%	1,05%	-0,43%	0,13%	2,47%
CLSIZE1	-6,60%	-3,77%	-3,63%	-6,07%	-3,66%	-6,34%
CLSIZE2	-8,40%	-2,21%	-0,43%	-0,24%	-0,87%	-2,53%
CLSIZE3	0,68%	4,09%	2,89%	2,31%	2,07%	5,75%
CLSIZE4	13,86%	5,80%	6,68%	7,00%	6,60%	4,30%
CLSIZE5	3,03%	-14,55%	-15,67%	-15,09%	-15,74%	-14,89%
CLSIZE6	1,21%	6,10%	3,60%	4,13%	4,24%	-0,16%
CLSIZE7	-22,46%	-0,56%	-1,69%	-1,15%	-2,11%	-1,56%
CLSIZE8	-38,45%	6,39%	6,27%	6,18%	5,26%	3,84%
CLSIZE9	-0,33%	-2,63%	-1,78%	-1,46%	0,24%	1,95%
STRATIO	-15,04%	0,92%	-2,90%	-4,25%	-5,54%	-0,29%

(*continued on next page*)

Table A1 (*continued*)

Variable	Balancing measures: standarised bias					
	Before matching	**NN1:1**	**NN1:3**	**NN1:5**	**NN1:10**	**NN1:10 C:0.15**
SCHRESOURCES1	-0,26%	1,89%	0,28%	-1,17%	-3,38%	-1,41%
SCHRESOURCES2	13,55%	5,48%	6,49%	5,86%	8,44%	3,86%
SCHRESOURCES3	14,05%	-2,88%	-3,19%	-2,70%	-1,25%	0,36%
SCHRESOURCES4	-26,62%	-2,93%	-2,83%	-2,10%	-4,97%	-3,29%
SCHSTAFF1	-7,17%	-4,22%	-5,89%	-6,87%	-7,90%	-3,72%
SCHSTAFF2	1,41%	-0,10%	-0,57%	-0,63%	0,39%	-1,09%
SCHSTAFF3	5,66%	-1,20%	-0,68%	1,34%	1,80%	0,37%
SCHSTAFF4	-1,95%	5,59%	7,15%	5,70%	4,77%	4,50%
PCRATIO	-2,31%	1,59%	1,12%	1,62%	3,59%	-1,22%
STANDARISED BIAS	**7,49%**					

Source: Own elaboration

Table A2
Medium socio-educational vulnerability bias.

Variable	Balancing measures: standarised bias					
	Before Matching	**NN1:1**	**NN1:3**	**NN1:5**	**NN1:10**	**NN1:10 C:0.15**
SVI	9,64%	-1,87%	-1,74%	-0,19%	0,80%	-5,19%
AGE	0,54%	-3,78%	-2,01%	-1,46%	-1,37%	-1,55%
SEX	-5,24%	0,00%	0,00%	0,00%	0,00%	0,00%
REPEAT	1,84%	0,00%	0,00%	0,00%	0,00%	0,00%
PREPRIMARY	-0,99%	0,37%	-1,08%	-0,16%	0,52%	-3,23%
CHILDHOOD	0,73%	-0,69%	1,25%	0,16%	-0,51%	4,20%
PRIMARY	1,03%	1,19%	-0,59%	-0,01%	-0,08%	-3,47%
IMMIG1	4,35%	0,00%	0,00%	0,00%	0,00%	0,00%
IMMIG2	3,25%	0,00%	0,00%	0,00%	0,00%	0,00%
NATIVE	-5,59%	0,00%	0,00%	0,00%	0,00%	0,00%
EDUFATHER1	2,60%	-3,23%	-0,35%	1,60%	1,21%	-2,83%
EDUFATHER2	-1,32%	-1,22%	-3,54%	-4,30%	-2,00%	-4,13%
EDUFATHER3	0,68%	-0,06%	-1,02%	-0,26%	-0,98%	-0,86%
EDUFATHER4	-2,14%	3,98%	3,77%	1,77%	1,16%	6,41%
EDUMOTHER1	4,19%	-1,58%	-0,38%	0,61%	1,67%	-3,03%
EDUMOTHER2	-0,02%	-4,07%	-2,14%	0,68%	1,30%	-0,61%
EDUMOTHER3	0,18%	-4,51%	-1,59%	-2,16%	-1,89%	-3,99%
EDUMOTHER4	-3,81%	7,96%	3,16%	0,73%	-0,86%	6,33%
DISCLIMA	-4,23%	-2,43%	-2,74%	-1,45%	0,18%	-3,22%
ESCS	-5,03%	1,06%	0,09%	-0,95%	-2,15%	3,48%
OWNDESKTOP	-4,61%	2,32%	-0,22%	-0,43%	-1,20%	5,82%
OWNBEDROOM	-2,44%	-0,54%	0,05%	-0,36%	0,09%	1,32%
OWNESPACE	-1,18%	-1,43%	1,33%	1,21%	0,09%	1,03%
PC	-0,67%	-4,06%	-0,28%	-0,30%	-0,69%	0,15%
INTERNET	-0,50%	-4,16%	-2,11%	0,71%	0,93%	-1,01%
BOOKS1	3,56%	-6,32%	-0,82%	-2,14%	0,12%	-4,27%
BOOKS2	2,33%	1,31%	-0,79%	-0,95%	-0,59%	-2,45%
BOOKS3	-2,14%	3,07%	1,20%	1,61%	1,50%	0,94%
BOOKS4	-2,36%	1,33%	1,27%	0,87%	-0,06%	-0,05%
BOOKS5	-2,36%	1,33%	1,27%	0,87%	-0,06%	-0,05%
BOOKS6	0,15%	-3,75%	-1,49%	-1,33%	-1,01%	2,30%
REGION	-15,69%	0,00%	0,00%	0,00%	0,00%	0,00%
CHARTER	-19,89%	7,17%	4,37%	2,04%	0,08%	7,75%
PRIVATE	1,15%	-1,75%	-0,92%	0,01%	-1,39%	4,41%
PUBLIC	17,58%	-5,57%	-3,48%	-1,88%	0,72%	-9,61%
MSIZE1	-6,94%	-7,23%	-6,49%	-4,97%	-2,38%	-5,86%
MSIZE2	2,21%	0,06%	0,68%	1,20%	1,56%	2,99%
MSIZE3	-3,16%	-2,50%	-3,35%	-2,37%	-3,35%	0,55%
MSIZE4	4,27%	7,13%	7,02%	5,23%	4,92%	2,03%
MSIZE5	-0,69%	-3,15%	-3,00%	-3,51%	-3,64%	-5,41%
CLSIZE1	0,53%	1,94%	1,13%	1,66%	2,34%	3,73%
CLSIZE2	-2,78%	1,17%	0,56%	0,46%	1,45%	-0,39%
CLSIZE3	-4,24%	-4,65%	-5,47%	-4,59%	-4,33%	-4,19%
CLSIZE4	11,58%	7,20%	8,03%	7,51%	7,59%	3,46%
CLSIZE5	5,16%	-4,69%	-5,63%	-5,60%	-6,96%	-4,12%
CLSIZE6	-4,85%	1,13%	-0,56%	-0,13%	0,20%	4,50%
CLSIZE7	-8,96%	-3,01%	-2,18%	-3,11%	-3,62%	-0,88%
CLSIZE8	-22,52%	6,92%	6,83%	6,54%	5,16%	0,78%
CLSIZE9	-3,83%	-4,19%	-1,98%	-2,42%	-2,61%	1,95%
STRATIO	-10,87%	6,37%	5,17%	3,87%	1,96%	4,18%
SCHRESOURCES1	2,74%	-4,10%	-2,49%	-2,70%	-3,15%	-5,79%
SCHRESOURCES2	8,22%	0,97%	2,39%	2,52%	2,68%	5,60%
SCHRESOURCES3	18,94%	1,40%	-0,39%	0,58%	1,33%	-2,10%
SCHRESOURCES4	-28,58%	-0,31%	-0,55%	-1,56%	-2,26%	-0,07%
SCHSTAFF1	-0,06%	-7,83%	-5,07%	-4,67%	-3,63%	-6,16%
SCHSTAFF2	0,16%	1,89%	-0,46%	0,24%	-1,17%	0,80%
SCHSTAFF3	4,15%	-1,75%	-1,37%	-1,76%	-0,11%	-2,28%

(*continued on next page*)

Table A2 (*continued*)

Variable	Balancing measures: standarised bias					
	Before Matching	**NN1:1**	**NN1:3**	**NN1:5**	**NN1:10**	**NN1:10 C:0.15**
SCHSTAFF4	-5,17%	7,67%	7,49%	6,60%	5,43%	8,03%
PCRATIO	-1,57%	-3,17%	-3,05%	-3,19%	-1,55%	-4,01%
STANDARISED BIAS	**5,02%**	**2,92%**	**2,14%**	**1,83%**	**1,67%**	**2,94%**

Source: Own elaboration

Table A3
Low socio-educational vulnerability bias.

Variable	Balancing measures: standarised bias					
	Befora Matching	**NN1:1**	**NN1:3**	**NN1:5**	**NN1:10**	**NN1:10 C:0.15**
SVI	7,90%	-2,98%	-0,56%	1,25%	1,94%	0,38%
AGE	1,91%	8,10%	5,08%	4,89%	5,74%	5,30%
SEX	3,90%	0,00%	0,00%	0,00%	0,00%	0,00%
REPEAT	1,10%	0,00%	0,00%	0,00%	0,00%	0,00%
PREPRIMARY	2,39%	-3,11%	-2,18%	-1,71%	-1,07%	-1,48%
CHILDHOOD	-2,79%	2,42%	2,46%	1,95%	1,42%	1,28%
PRIMARY	1,30%	2,60%	-0,87%	-0,75%	-1,15%	0,79%
IMMIG1	5,04%	0,00%	0,00%	0,00%	0,00%	0,00%
IMMIG2	3,30%	0,00%	0,00%	0,00%	0,00%	0,00%
NATIVE	-6,16%	0,00%	0,00%	0,00%	0,00%	0,00%
EDUFATHER1	1,56%	-3,68%	-1,12%	1,46%	0,21%	0,65%
EDUFATHER2	1,14%	1,36%	2,65%	0,83%	2,71%	1,19%
EDUFATHER3	1,07%	5,02%	3,27%	3,13%	1,42%	2,39%
EDUFATHER4	-3,28%	-2,03%	-3,85%	-4,78%	-3,50%	-3,63%
EDUMOTHER1	2,20%	-7,48%	-0,75%	0,97%	0,51%	3,42%
EDUMOTHER2	2,34%	2,47%	4,57%	2,88%	2,55%	2,23%
EDUMOTHER3	0,69%	4,72%	-0,73%	-0,50%	-1,03%	-2,42%
EDUMOTHER4	-4,20%	0,69%	-2,23%	-2,60%	-1,52%	-2,63%
DISCLIMA	-6,82%	0,03%	2,40%	1,30%	-1,36%	4,10%
ESCS	-3,49%	1,86%	-0,75%	-2,34%	-1,39%	-1,90%
OWNDESKTOP	0,63%	3,10%	3,23%	3,77%	2,41%	1,43%
OWNBEDROOM	-1,07%	-0,76%	1,70%	0,90%	-0,67%	-0,90%
OWNESPACE	-1,33%	3,29%	2,29%	1,15%	0,51%	2,12%
PC	-3,13%	1,31%	0,06%	-0,80%	-1,09%	1,19%
INTERNET	-3,84%	-2,14%	-1,78%	-2,16%	-2,67%	0,37%
BOOKS1	-0,95%	-1,98%	-1,53%	-0,80%	-1,44%	1,68%
BOOKS2	8,35%	-9,75%	-3,12%	1,32%	4,27%	-10,81%
BOOKS3	-2,69%	0,42%	0,43%	-1,89%	-2,13%	2,66%
BOOKS4	-1,49%	4,04%	3,32%	2,39%	1,66%	3,78%
BOOKS5	-1,49%	4,04%	3,32%	2,39%	1,66%	3,78%
BOOKS6	-2,40%	5,22%	2,14%	1,83%	1,61%	1,35%
REGION	-19,07%	0,00%	0,00%	0,00%	0,00%	0,00%
CHARTER	-12,37%	-2,43%	-1,84%	-1,27%	-2,04%	-0,51%
PRIVATE	1,60%	-2,94%	-2,91%	-3,82%	-4,64%	-1,28%
PUBLIC	10,59%	3,90%	3,33%	3,31%	4,49%	1,19%
MSIZE1	-9,00%	-2,04%	-2,76%	-3,09%	-5,50%	-3,61%
MSIZE2	2,87%	-2,56%	-4,30%	-4,50%	-3,93%	-5,08%
MSIZE3	-0,41%	1,63%	2,10%	2,31%	3,26%	4,57%
MSIZE4	-2,36%	5,53%	6,16%	5,90%	5,42%	5,01%
MSIZE5	5,85%	-5,95%	-4,63%	-4,06%	-4,16%	-5,18%
CLSIZE1	-1,87%	-3,49%	-5,81%	-8,18%	-8,70%	-4,53%
CLSIZE2	4,26%	0,76%	-0,38%	-0,44%	0,08%	0,21%
CLSIZE3	-3,04%	-9,04%	-5,78%	-3,96%	-4,64%	-4,75%
CLSIZE4	3,31%	8,55%	6,25%	5,69%	6,36%	8,62%
CLSIZE5	13,06%	1,71%	5,16%	5,01%	7,11%	0,73%
CLSIZE6	-10,04%	-1,47%	-2,33%	-2,51%	-3,40%	-1,95%
CLSIZE7	-18,10%	1,28%	0,85%	0,26%	-0,64%	-1,97%
CLSIZE8	-8,08%	6,79%	5,59%	4,59%	4,83%	1,49%
CLSIZE9	-12,41%	-4,14%	-6,63%	-6,88%	-9,73%	-6,67%
STRATIO	-1,64%	1,94%	3,45%	4,40%	3,30%	-2,93%
SCHRESOURCES1	-2,15%	-0,83%	0,48%	2,18%	2,92%	-1,68%
SCHRESOURCES2	3,26%	-1,71%	-2,84%	-2,57%	-2,02%	-1,38%
SCHRESOURCES3	16,80%	3,60%	3,95%	4,61%	5,07%	-0,54%
SCHRESOURCES4	-18,38%	-1,79%	-1,84%	-3,53%	-4,81%	2,47%
SCHSTAFF1	5,57%	1,15%	2,72%	2,87%	4,78%	-0,33%
SCHSTAFF2	-2,74%	-1,89%	-4,08%	-1,64%	-2,11%	0,12%
SCHSTAFF3	1,53%	1,07%	-0,75%	-1,73%	-1,97%	-3,08%
SCHSTAFF4	-3,99%	0,00%	3,48%	1,31%	0,27%	3,94%
PCRATIO	-8,10%	-3,57%	-4,76%	-4,49%	-3,08%	-6,14%
STANDARISED BIAS	**4,96%**	**2,82%**	**2,57%**	**2,47%**	**2,66%**	**2,44%**

Source: Own elaboration

Table A4
Detailed results with means, standard deviation, and significance tests.

Knowledge Area	Vulnerability	Mean			Standard deviation		Student T-test
		Treatment	Control	Difference	Treatment	Control	
Global	**High**	524,4203	529,0756	-0,88%	87,9150	78,1914	-2423**
	Medium	526,1008	527,0534	-0,18%	85,4509	83,7104	-0,8014
	Low	527,7034	529,0908	-0,26%	89,6892	85,136	-0,9062
Science	**High**	494,1336	497,4713	-0,67%	80,513	78,1914	-2,6590***
	Medium	495,1421	495,1739	-0,01%	76,9595	76,7394	-0,6682
	Low	496,5418	496,3673	0,04%	82,4357	78,8132	-0,2551
Reading	**High**	486,9270	490,9888	-0,83%	87,4486	85,2437	-2,3861**
	Medium	487,0349	488,8584	-0,37%	84,6129	85,4509	-1,3902
	Low	487,7277	492,4856	-0,97%	89,6892	84,136	-2,1836**
Mathematics	**High**	497,5572	500,0567	-0,50%	79,0592	74,8402	-2,0783**
	Medium	497,6374	496,7765	0,17%	75,1958	74,8842	0,3998
	Low	498,6908	499,0500	-0,07%	89,5627	77,615	-0,4032

Source: Own elaboration

References

Agasisti, T., Longobardi, S., 2014. Inequality in education: can Italian disadvantaged students close the gap? J. Behav. Exp. Econ. 52, 8–20. https://doi.org/10.1016/j.socec.2014.05.002.

Aisenson, G., Legaspi, L., Valenzuela, V., Duro, L., De Marco, M., Moulia, L., Del Re, V., Bailac, K., Suescún, J., 2011. Las anticipaciones de futuro de jóvenes en situación de vulnerabilidad socio-educativa. Anu. De. Invest. XVIII, 143–152.

Alegre, M.Á., Subirats, J., 2013. Sistemas y políticas educativas comparadas: transformaciones, convergencias y divergencias en los países europeos. In: Del Pino, En.E., Rubio, M.J. (Eds.), Los Estados de Bienestar en la Encrucijada, 2ª ed. Tecnos, pp. 342–369.

Álvarez, M., González, M., Harrington, A., Peláez, E., Ribotta, B., 2004. Vulnerabilidad socio-educativa. Un análisis transversal de la realidad de Córdoba. Comunic-arte, Córdoba.

Aramendi, P., Vega, A., Buján, K., 2012. Los programas de diversificación curricular y los programas de cualificación profesional inicial ¿una alternativa al fracaso escolar? Rev. Española De Pedagogía 252, 237–256.

Barberá, E., Candela, C., Ramos, A., 2008. Elección de carrera, desarrollo profesional y estereotipos de género. Rev. De. Psicol. Soc. 23 (2), 275–285. https://doi.org/10.1174/021347408784135805.

Becker, M., Neumann, M., Tetzner, J., Böse, S., Knoppick, H., Maaz, K., Baumert, J., Lehmann, R., 2014. Is early ability grouping good for high-achieving students' psychosocial development? Effects of the transition into academically selective schools. J. Educ. Psychol. *106* (2), 555–568. https://doi.org/10.1037/a0035425.

Benito, R., González-Balletbò, I., 2012. ¿Puede el efecto composición reducir las desigualdades educativas? Una mirada sobre los efectos de la segregación escolar en España y los países del modelo de integración uniforme. Rev. De. la Asoc. De. Sociol. ía De. la Educ. óN. VI (1), 49–71.

Blanco-Varela, B., 2022. Unha análise socioeconómica da vulnerabilidade en Galicia: o camiño dende a escola até a inserción laboral [Unpublished PhD Thesis]. Universidade de Santiago de Compostela.

Bonal, X., Bellei, C., 2018. The renaissance of school segregation in a context of globalization. In: Bonal, X., Bellei, C., Bonal, X., Bellei, C. (Eds.), Understanding school segregation. Patterns, causes and consequences of spatial inequalities in education. Bloomsbury Academic, pp. 1–35.

Busemeyer, M.R., Nikolai, R., 2010. Education. In: Castles, Francis G., Lewis, J., Obinger, H., Pierson, C. (Eds.), The Oxford Handbook of the Welfare State. Oxford University Press, pp. 494–510. https://doi.org/10.1093/oxfordhb/9780199579396.001.0001.

Cavaco, C., Alves, N., Guimarães, P., Feliciano, P., Paulos, C., 2021. Teachers' perceptions of school failure and dropout from a gender perspective: (re)production of stereotypes in school. Educ. Res. Policy Pract. 20 (29–44) https://doi.org/10.1007/s10671-020-09265-7.

den Butter, F.A.G., van den Eyden, J.A.C., 1998. A pilot index for environmental policy in the Netherlands. Energy Policy 26 (2), 95–101. https://doi.org/10.1016/S0301-4215(97)00010-4.

Caballé, J., 2016. Intergenerational mobility: measurement and the role of borrowing constraints and inherited tastes. Ser. J. Span. Econ. Assoc. 7, 393–420. https://doi.org/10.1007/s13209-016-0149-2. ⟨https://doi:10.1007/s13209-016-0149-2⟩.

Campbell, T., 2021. In-class 'ability'-grouping, teacher judgements and children's mathematics self-concept: evidence from primary-aged girls and boys in the UK Millennium Cohort Study. Camb. J. Educ. *51* (5), 563–587. https://doi.org/10.1080/0305764X.2021.1877619.

Catalayud, M.A., 2018. Los agrupamientos escolares a debate. Tend. Pedagógicas (32), 5–14. https://doi.org/10.15366/tp2018.32.001.

Dumay, X., Durpiez, V., 2008. Does the school composition effect matter? Evidence from Belgian data. Br. J. Educ. Stud. *54* (4), 440–477. https://doi.org/10.1111/j.1467-8527.2008.00418.x.

Dupriez, V., Dumay, X., Vause, A., 2008. How do school systems manage pupils heterogeneity? Comp. Educ. Rev. *LII* (2), 245–273. https://doi.org/10.1086/528764.

Dupriez, V., Monseur, C., Campenhoudt, M.V., Lafontaine, D., 2012. Social inequalities of post-secondary educational aspirations: influence of social background, school composition and institutional context. Eur. Educ. Res. J. *11* (IV), 504–519. https://doi.org/10.2304/eerj.2012.11.4.504.

Ebert, U., Welsh, H., 2004. Meaningful environmental indices: a social choice approach. J. Environ. Econ. Manag. 47 (2), 270–283. https://doi.org/10.1016/j.jeem.2003.09.001.

Fernández, C., García, O., Galindo, E., 2017. Escuela o barbarie. Entre el neoliberalismo salvaje y el delirio de la izquierda, 3ª ed. Akal.

Ficano, C.C., 2012. Peer effects in college academic outcomes–gender matters! Econ. Educ. Rev. *31*, 1102–1115. https://doi.org/10.1016/j.econedurev.2012.07.012.

Francis, B., Archer, L., Hodgen, J., Pepper, D., Taylor, B., Travers, M.C., 2015. Exploring the relative lack of impact of research on 'ability grouping' in England: a discourse analytic account. Camb. J. Educ. *47*, 1–17. https://doi.org/10.1080/0305764X.2015.1093095.

Francis, B., Craig, N., Hodgen, J., Taylor, B., Tereshchenko, A., Connolly, P., Archer, L., 2020. The impact of tracking by attainment on pupil self-confidence over time: demonstrating the accumulative impact of self-fulfilling prophecy. Br. J. Sociol. Educ. *41* (5), 626–642. https://doi.org/10.1080/01425692.2020.1763162.

Gabarró, D., 2010. ¿Fracaso escolar? La solución inesperada del género y la coeducación. Boira Editorial.

Garrido-Yserte, R., Gallo-Rivera, M.T., Martínez-Gautier, D., 2019. Más allá de las aulas: los determinantes del bajo rendimiento educativo en España y el fracaso de las políticas públicas. International Review of. Econ. Policy 1 (1), 86–106. https://doi.org/10.7203/IREP.1.1.16459.

Gracia, P., Esping-Andersen, G., 2015. Fathers' child care time and mothers' paid work: a cross-national study of Denmark, Spain, and the United Kingdom. Fam. Sci. *6* (1), 277–281. https://doi.org/10.1080/19424620.2015.1082336.

Guill, K., Lüdtke, O., Köller, O., 2017. Is academic tracking related to gains in learning competence? combining propensity score matching and analysis of differential item functioning. Learn. Instr. 47, 43–52. https://doi.org/10.1016/j.learninstruc.2019.101286.

Hermans, E., van den Bossche, F., & Wets, G. 2007. Impact of Methodological Choices on Road Safety Ranking. SAMO 2007, 5th International Conference on Sensitivity Analysis of Model Output, Budapest (Hungary), June 18-22.

INEE, Instituto Nacional de Evaluación Educativa, 2016. Alumnos de bajo rendimiento: por qué se quedan atrás y cómo se les puede ayudar. In: Boletín de educación, 47. ⟨https://www.educacionyfp.gob.es/dctm/inee/boletines/educainee47.pdf?documentId=0901e72b8202aa4e⟩.

Kang, S.M., 2002. A sensitivity analysis of the Korean composite environmental index. Ecol. Econ. 43 (2/3), 159–174. https://doi.org/10.1016/S0921-8009(02)00207-0.

Martínez-Garrido, C., Lange, R., 2019. Rendimiento acedémico y miedo. Un estudio exploratorio. Form. óN. De. Profr., Invest. óN. Y. Justicia Soc., Colab. óN. *23* (4), 280–297. https://doi.org/10.30827/profesorado.v23i4.11722.

Meier, V., Schütz, G., 2007. The economics of tracking and non-tracking. Ifo Working Paper Series 50. Ifo Institute/Leibniz Institute for Economic Research at the University of Munich.

Munda, G., 2005. Measuring sustainability: a multi-criterion framework. Environ. Dev. Sustain. 7, 117–134. https://doi.org/10.1007/s10668-003-4713-0.

Narciso, L. 2022. Els agrupaments escolars: impactes sobre les oportunitats educatives dels i les joves. En A. Tarabini (Ed.), Els reptes de l'educació a Catalunya 1ª ed.,173-216.

OECD [The Organization for Economic Cooperation and Development]. (2019). PISA 2018 Database. ⟨https://www.oecd.org/pisa/data/2018database/⟩.

OECD, The Organization for Economic Cooperation and Development, 2020. PISA 2018 Results (Volume V): Effective Policies, Successful Schools. OECD Publishing, Paris.

OECD/European Union/EC-JRC, 2008. Handbook on constructing composite indicators. Methodology and user guide. OECD Publishing, Paris. https://doi.org/10.1787/9789264043466-en.

Owens, A., 2017. Income segregation between school districts and inequality in students' achievement. Sociol. Educ. XX (X), 1–27. https://doi.org/10.1177/0038040717741180.

Pekkarinen, T., 2018. School tracking and intergenerational social mobility. IZA World Labor 56. https://doi.org/10.15185/izawol.56.v2.
Rogero-García, J., Andrés-Candelas, M., 2014. Gasto público y de las familias en educación en España: diferencias entre centros públicos y concertados. Rev. Española De. Invest. Socio ógicas (*147*), 121–132. https://doi.org/10.5477/cis/reis.147.121.
Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika *70* (1), 41–55. https://doi.org/10.1093/biomet/70.1.41.
Saisana, M., Tarantola, S., 2002. State - of - the - Art Report on Current Methodologies and Practices for Composite Indicator Development. Joint Research Centre, European Commission.
Sicilia, G., Simancas, R., 2018. Equidad Educativa En España: Comparación Regional A Partir De Pisa 2015. Fundación Ramón Areces.
Stuart, E.A., 2010. Matching methods for causal inference: a review and a look forward. Stat. Sci. Rev. J. Inst. Math. Stat. *25* (1), 1 https://doi.org/10.1214%2F09-STS313.
Stuart, E.A., King, G., Imai, K., Ho, D., 2011. MatchIt: nonparametric preprocessing for parametric causal inference. J. Stat. Softw. *42* (8) https://doi.org/10.18637/jss.v042.i08.
Tan, C.Y., Dimmonck, C., 2020. The relationships among between-class ability grouping, teaching practices, and mathematics achievement: a large-scale empirical analysis. Educ. Stud. *48* (4), 471–489. https://doi.org/10.1080/03055698.2020.1780416.
Taylor, B., Francis, B., Archer, L., Hodgen, J., Pepper, D., Tereshchenko, A., Travers, M.-C., 2017. Factors deterring schools from mixed attainment teaching practice. Pedagog. Cult. Soc. 25 (3), 327–345. https://doi.org/10.1080/03055698.2020.1780416.
Tsaur, S.-H., Lin, Y.-C., J.-H, 2006. Evaluating ecotourism sustainability from the integrated perspective of resource, community and tourism. Tour. Manag. 27 (4), 640–653. https://doi.org/10.1016/j.tourman.2005.02.006.
Uysal, S., Banoglum, K., 2018. Hogging the middle lane: how student performance heterogeneity leads Turkish schools to fail in PISA? Cypriot J. Educ. Sci. 13 (2), 448–460. https://doi.org/10.18844/cjes.v13i2.3196.
Zhang, Y., Chen, D., Wang, W., 2014. The heterogeneous effects of ability grouping on national college entrance exam performance – evidence from a large city in China. Int. J. Educ. Dev. 39, 80–91. https://doi.org/10.1016/j.ijedudev.2014.08.012.