

OECD Economics Department Working Papers No. 1460

Public spending in education and student's performance in Colombia

Laura Heras Recuero,

Eduardo Olaberría

https://dx.doi.org/10.1787/282d9700-en

OECD

Unclassified **ECO/WKP(2018)7**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **19-Feb-2018**

English - Or. English

ECONOMICS DEPARTMENT

ECO/WKP(2018)7
Unclassified

PUBLIC SPENDING IN EDUCATION AND STUDENTS' PERFORMANCE IN COLOMBIA

ECONOMICS DEPARTMENT WORKING PAPERS No. 1460

By Laura Heras Recuero and Eduardo Olaberría

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Authorised for publication by Alvaro Pereira, Director, Acting Chief Economist, Economics Department.

All Economics Department Working Papers are available at www.oecd.org/eco/workingpapers

JT03426965

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the author(s).

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works.

Comments on Working Papers are welcomed, and may be sent to OECD Economics Department, 2 rue André-Pascal, 75775 Paris Cedex 16, France, or by e-mail to eco.contact@oecd.org.

All Economics Department Working Papers are available at www.oecd.org/eco/workingpapers.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

© OECD (2018)

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for commercial use and translation rights should be submitted to **rights@oecd.org**

ABSTRACT/RÉSUMÉ

Public spending in education and student's performance in Colombia

This paper investigates if higher public spending in education and better teacher qualifications are related to student's performance, using data from Saber 11, a national standardized test conducted by *Instituto Colombiano para la Evaluación de la Educación*. The estimation exploits differences in both policy variables across regions and employs interactions to study if more investment in public education and higher teacher qualifications can help increase average performance and reduce the impact that socioeconomic factors, such as family income, have on student performance. The analysis proposes a model where student performance in Mathematics and Language are dependent not only on the variables of interest of this paper, but also on economic, social and cultural status, sex and age of students, and school characteristics. The results show that students' characteristics and their environment, school features and departmental differences in the policy variables explain roughly 20% of the variation in education performance in Colombia, a relatively high percentage when compared to those found by other studies focusing on OECD countries and based on PISA. After controlling for students' and school characteristics, the results show that in Colombia, public spending per student and teacher qualifications are positively related to better learning outcomes. For the first one, the results suggest that if all regions reach the level of spending per student of Bogota – the region with the highest spending – average math scores can increase by 3.8 to 4.3 points (around 8%), depending on the regions, with the highest improvement for low income students.

This Working Paper relates to the OECD Economic Survey of Colombia (www.oecd.org/eco/surveys/economic-survey-colombia.htm)

JEL classification codes: H0, I20, I24, J00, J1.
Keywords: Quality of education, language scores, maths scores, human capital, public spending, inequality

****

Les dépenses publiques liées à l'éducation et la performance des étudiants en Colombie

Cet article étudie si l'augmentation des dépenses publiques dans l'éducation et de meilleures qualifications des enseignants sont liées à la performance des élèves, d'après les données de Sabre 11, un test national standardisé mené par *l'Instituto Colombiano para la Evaluación de la Educación*. L'estimation exploite les différences entre les deux politiques variables, à travers les régions et emploie des interactions afin d' étudier si un meilleur investissement dans l'éducation publique et des qualifications d'enseignement supérieures peuvent aider à augmenter la performance moyenne et réduire l'impact des facteurs socioéconomiques, tels que les revenus par famille. L'analyse propose un modèle où le niveau des élèves en mathématiques et en langues dépendent non seulement des variables d'intérêt de ce document, mais aussi de l'état économique, social et culturel, du sexe et de l'âge des élèves et des caractéristiques scolaires comprises comme des variables indépendantes . Les résultats montrent que les caractéristiques des étudiants et leur environnement, les caractéristiques de l'école et les différences départementales dans les variables politiques, représentent environ 20% de la variation du rendement scolaire en Colombie; un pourcentage relativement élevé, comparé à d'autres études portant sur les pays de l'OCDE et basé sur l'étude PISA. Après avoir vérifié les caractéristiques des élèves et de l'école, les résultats montrent qu'en Colombie, les dépenses publiques par étudiant et les qualifications des enseignants sont positivement liées à l'amélioration du niveau d'apprentissage. Dans le premier cas, les résultats suggèrent que si toutes les régions atteignent le niveau de dépenses par élève de Bogota - la région ayant les dépenses les plus élevées - les résultats moyens en mathématiques pourraient augmenter de 3,8 à 4,3 points (environ 8%) selon les régions, avec une hausse plus importante pour les étudiants à faibles revenus.

JEL classification codes: H0, I20, I24, J00, J1.
Mots clés : Qualité d' éducation, niveau en langues, niveau en mathématiques, capital humain, dépenses publiques, inégalité

TABLE OF CONTENTS

PUBLIC SPENDING IN EDUCATION AND STUDENT'S PERFORMANCE IN COLOMBIA 5

Introduction.....5
Review of the literature.....6
Stylized facts on student scores and public spending in education in Colombia.....8
Data.....11
Empirical strategy.....18
Results: the determinants of student's performance.....19
Interaction terms: measuring the role of public spending on education and teacher qualifications.....22
Probit estimation: determinants of low-performing in Saber 11.....27
Summary discussion and Conclusions.....30

REFERENCES.....32

STATISTICAL APPENDIX.....34

Tables

1. Baseline estimation.....20
2. Baseline estimation with interactions for Mathematics.....23
3. Baseline estimation with interactions for Language.....24
4. Marginal effects.....30

Figures

1. PISA results in Science.....8
2. Student and teacher behavior hindering learning (PISA 2015).....9
3. Annual expenditure per student by educational level (2011).....10
4. Distribution across the education system of responsibility for school resources.....10
5. Average score in Mathematics and Language by department.....13
6. Distribution of test takers by household income.....14
7. Average score by monthly household income and department.....15
8. Public funding per student and academic performance.....16
9. Percentage of teachers holding postgraduate diploma.....18
10. Impact of public funding on student's performance.....26

PUBLIC SPENDING IN EDUCATION AND STUDENT'S PERFORMANCE IN COLOMBIA

By Laura Heras Recuero and Eduardo Olaberría[1]

Introduction

1. Like many other emerging economies, Colombia has made great progress over the last two decades increasing student enrolment rates (OECD, 2016a). Moreover, it has done so alongside improving learning outcomes. However, the efforts made by emerging countries to increase enrolment have been met with a challenge of quality (Pritchett, 2006; Carlitz, 2009). Low performance in education is a major concern, as it carries high economic and social costs (Hanushek and Woessmann, 2015). As a result, research is increasingly focusing on the quality aspect of the education being offered in the schools (OECD, 2013; 2016d), and the policies that could help achieve basic literacy and numeracy. Indeed, this is one important goal of the SDGs (OECD, 2016b). One of the policies that could potentially help reach this goal is the level of public spending in education, as adequate education funding is a necessary condition for the provision of quality education. Understanding the role of public spending in education as determinant of school quality and student outcomes is, therefore, of first-order significance.

2. This paper tackles the question of whether and how public spending in education is related to student outcomes using evidence from Colombia. Colombian regions vary significantly in school quality and in public spending per student in education (OECD, 2016a; Bonilla and Galvis, 2014; Duarte et al., 2012). Therefore, it provides a useful setup to study the relationship between higher public spending in education and student performance. The objectives of this paper are twofold: it aims to (1) estimate how differences in district per-student spending are related to average students' outcomes in math and reading, and (2) shed light on underlying mechanisms through which public spending may affect students'outcomes by considering the role of student characteristics, e.g., family income.

3. To do this, this paper models student's performance in Colombia's Saber 11 – a national standardized test conducted by *Instituto Colombiano para la Evaluación de la Educación*. Saber 11 is mandatory for all senior year high school students who wish to access higher education. However, each tertiary education institution establishes its own score criteria for admission. The estimation exploits differences in policy indicators across regions, and employs interactions to study if more investment in public education can help reduce the impact that socioeconomic factors, such as family income and level of education of the mother, have on student performance. The analysis proposes a model where student performance in mathematics is the dependent variable, and economic, social and cultural status of students and schools, sex and age are included as independent variables.

4. The results show that students' characteristics and their environment (i.e. sex, age and economic, social and cultural status of students and schools, level of education of their mothers) explains roughly 20% of the variation in education performance in Colombia, a relatively high percentage when comparing with other studies focusing on OECD countries and based on PISA (Avendaño et al., 2016). This finding can be attibuted to larger inequalities in Colombia, reflected by a Gini index before taxes and transfers of 53.5[1], compared to 40.0 for OECD countries in 2014. After controlling for students' characteristics and

[1]Laura Heras Recuero was a Consultant for the OECD Economics Department and Eduardo Olaberría was Senior Economist and Latin American Coordinator at the OECD Economics Department at the time of writing. The authors would like to thank Anna Pons from the Directorate for Education and Skills (EDU), Jens Arnold, Christine de la Maisonneuve and Piritta Sorsa for helpful comments and suggestions, and Anthony Bolton and Assa Fofana for editorial assistance (all from the Economics Department).

their environment, the results show that in Colombia, public spending per student is positively related to better learning outcomes. In particular, the results suggest that if all regions would reach the level of spending per student of Bogota – the region with the highest spending – average math scores could increase by 3.8 to 4.3 points (around 8%), depending on the regions, with the highest improvement for low income students and poorest regions.

5. The rest of the paper is organized as follows. Section two reviews the literature on education quality putting emphasis on the measurement. Section three discusses key stylized facts of student's performance and public spending in education in Colombia. Sections four and five present, respectively, the database and the methodology used. Section six discusses the main results. Finally, gives a summary and provides conclusions of the study.

Review of the literature

6. The literature covering the determinants of student performance is extensive and can be divided in three areas of analysis. Some authors focus on the individual or household level, analyzing the role played by thestudent personal characteristics and her socioeconomic background on determining academic outcomes. The second level refers to school characteristics. Some factors covered at this level are school infrastructure and facilities, teacher's education, organizational features of the schools and so on. The third levelcovers the institutional level; i.e., the features of the educational system and therefore the role played by public spending on education.

7. Regarding the first level, student's characteristics and her socioeconomic background, the work of Coleman et al. (1966) is considered the first precedent. With a focus on the US, he shows the student's family background is the main determinant of student performance. Using a panel of 58 countries, Lee & Barro (2001) find similar results, concluding that family characteristics, particularly family income and years of education of parents, have a strong relation with student performance.

8. In the case of Colombia, OECD (2016a) highlights socioeconomic background determines educational outcomes in Colombia far more than in most OECD countries. This has repeatedly been proved by the literature with a focus on the country. Gaviria and Barrientos (2001) analyzed the outcomes of 11st grade students who took the ICFES national test in Bogotá in 1999 using a sample of 50,000 students in 1,300 schools. They find socioeconomic background of students and school characteristics determine around 25-30 percent of academic results at ICFES test, a higher proportion than in developed countries such as the US. However, when considering only those students attending the same school, the authors find the socioeconomic variables are not relevant for explaining academic outcomes. They conclude the effect of socioeconomic variables is transmitted by the decision of parents of enrolling their children in different quality schools. This is high-socioeconomic level students will be enrolled in high performance schools and low-socioeconomic level students will do in low quality schools. The same result is also highlighted by Sarmiento, Becerra & González (2000) who also study the determinants of academic performance in Colombia.

9. More recently, Duarte, Bos & Moreno (2012) deepen the understanding of academic performance determinants using the national results of ICFES national test, called Saber, in 2009 for students of grades 5^{th} (Saber 5) and 9^{th} (Saber 9). They conclude score differences can be explained mainly in terms of school differences which show a high level of segregation of students in Colombia. Therefore authors highlight students face unequal learning opportunities depending on their socioeconomic background since those students with high socioeconomic level will go to high-performing schools.

[1] Source: World Bank (http://data.worldbank.org/indicator/SI.POV.GINI).

10. Using results from Saber from 2011, García et al. (2013) conclude the individual level explains the 70% of differences in Saber results, the 23% is explained at the school level and the rest, 8%, is explained at the municipality level. At individual level, they find the main explanatory variables are household socioeconomic level and parenting education. At school level, an important finding to be mentioned is that when controlling by schooling fees and socioeconomic level of students, they find private schools underperform in Saber test. This leads the authors to conclude that attending to a private school does not guarantee good academic performance but attending a high-fees institution.

11. School characteristics are also a very important factor explaining student's performance in Colombia. Gaviria and Barrientos (2001) use evidence from schools in Bogotá and show that school features, particularly school infrastructure and teacher's qualifications influence academic outcomes more than individual features, but only in the case of private schools. For public schools, these variables do not show impact on academic outcomes. Therefore authors argue its impact must be mediated by other variables, such as teacher incentives or the organizational structure of the school. However, OECD (2013a) shows material infrastructure is particularly relevant as a determinant of academic performance in both public and private schools in those countries with a GDP per capita under USD 20,000. This contrast can be explained, in the case of Colombia, by different dymanics in Bogotá from the rest of the country due to the capital's higher economic development and public funding in education.

12. In line with this result, Bonilla & Galvis (2014) analyze the effect of the teacher's qualifications on academic outcomes in Saber test for 11st grade students in 2009. They conclude that teacher's qualifications, in particular the % of teachers holding a postgraduate diploma, have a positive and significant effect on academic outcomes even when controlling by individual characteristics, socioeconomic background and school and municipality features. However, the effect of teacher's qualification is higher in Mathematics than in Language. Aiming to broad their analysis and to avoid a potential endogeneity problem between the individual and school characteristics, they run liner regressions using instrumental variables and obtain similar results for teacher's qualifications. But when they study separately public and private schools, they find that while teachers in public schools have higher qualifications, their impact is lower than in the case of private schools. Therefore, they authors suggest other variables related to additional resources to the educational process might be playing an important role for strengthening the potential of teacher's qualifications.

13. Finally, some authors have analyzed the relationship between school funding and school resources and student's performance. However, the results are not conclusive. This is in part due to the difficulties for estimating this relation, as there are organisational processes that are not easy to compute and so there is little empirical evidence at this respect (Hanushek, 2004). Hanushek (2004) highlights that the current evidence between school resources and academic performance is not conclusive and results depend on the interaction between monetary resources, teacher's quality and so on. OECD (2013a) also points out that increasing funding in education only will generate better academic performance when resources are allocated in an effective way. Good governance of resources and well-designed policies seems to play an importante role as it has been shown in the Brazilian case. Incentives mechanisms such as a performance-based pay for teachers and municipals as well as resource management at municipal level has helped to improve academic performance in the country (OECD, 2015).

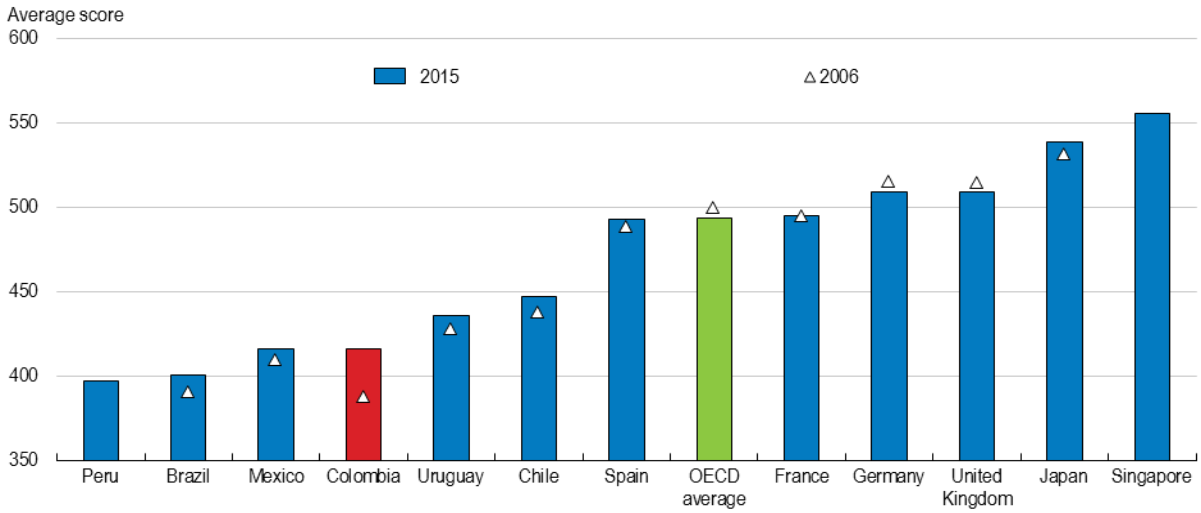
14. With a focus on Colombia, Gaviria and Barrientos (2001) show the substantial increase of public funding in education implemented in last decade of the Twentieth Century in Colombia did not help to close the gap in academic performance between public and private schools which remains constant. The authors highlight the importance of strengthening incentives to teachers and improving the organizational structure of the schools together with the increase of public funding in order to ensure impact on academic performance and foster equal opportunity to access quality education for all.

15. In this paper, we contribute to the existing literature demonstrating empirically that increases on public spending in education and teacher qualifications at regional level has a positive impact onstudent's academic performance and therefore are particularly beneficial to those individuals coming from low socioeconomic backgrounds.

Stylized facts on student scores and public spending in education in Colombia

16. Colombia is putting quality education in the forefront. Recent policy changes (OECD, 2016a) have been somewhat successful at improving quality. Evidence from PISA 2015 shows that mean performance in Colombia has slightly increased compared with previous years when taking into consideration PISA changes made to the test design, administration and scaling. For example, in science, performance seems to have increased 28 score points since 2006, the second largest improvement among the 52 education systems with comparable data, but it actually increased only 3 score point when considering the effect of PISA changes, particularly the introduction of a new scaling approach (OECD, 2016c). This could also be the case for reading and mathematics scores, which could be showing skewed improvements when compared with previous results (improvement of 40 score points compared to 2006 for reading and 20 score points for math).

Figure 1. **PISA results in Science**



Note: The graph shows PISA results without adjusting for changes made to test design between 2006 and 2015.

Source: OECD (2016c).

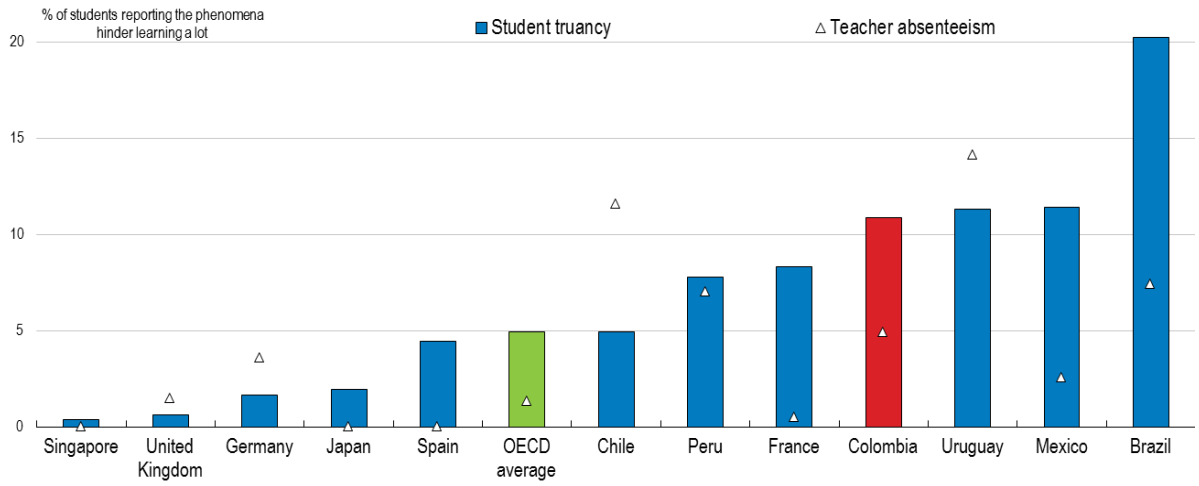
17. However, quality of education in Colombia remains relatively low. Students score 416 points in science, on average, below the OECD average and the mean score of Chile (447 points), at the level of Costa Rica, Georgia, Mexico, Montenegro, Qatar and Thailand, and above that of Brazil (410 points) and Peru (397 points) (see Figure 1). The percentage of students in Colombia who had repeated a grade is the second largest among all participants in PISA 2015. Moreover, 66% of students are low achievers in mathematics.

18. Studies done in the field of basic education have revealed that many pupils are still leaving school without mastering a minimum set of cognitive and non-cognitive skills (Chimombo, 2005). Also, the teaching staff not always meets the minimum standards of her profession (Bruns & Luque, 2015). Poorly qualified teachers and ill-equipped schools with scant or completely no learning materials remain a common scenario in many regions of the country. Yet achieving quality education for all fundamentally

relies on assuring a proper learning environment at school. It becomes pertinent therefore to identify the main challenges and factors that affect quality of education in Colombia. It would also include to identify the required policy directions around the challenge of quality so that there is no trade-off between quality and quantity in the process. This paper aims to contribute to it potentiallysheding some light on the factors that determine quality of education in Colombia at student, school and regional level.

19. In Colombia, 26% of students attend schools whose principal considers that the capacity to provide instruction is hampered a lot by the lack of educational material, significantly higher share than the OECD average (6%) or Chile (1%) (OECD, 2016d). 11% of students in Colombia attend schools whose principal reported that student truancy hinders learning a lot, compared to 5% of students across OECD countries who attend such schools. About 5% of students in Colombia are in schools whose principal reported that teacher absenteeism hinders learning a lot, compared to 1% of students across OECD countries who attend such schools (Figure 2).

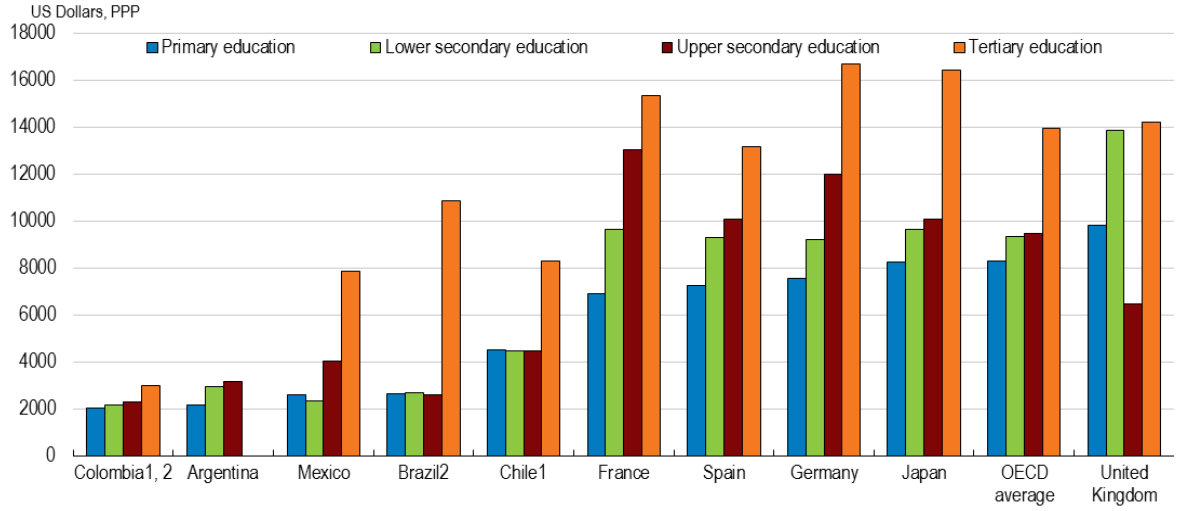
Figure 2. **Student and teacher behavior hinders learning (PISA 2015)**



Source: OECD (2016d).

20. Education is one of the largest components of government spending in the country, amassing 6.7% of GDP, slightly above OECD average of 6.1% (OECD, 2014). However, public spending per student in Colombia is far below the OECD average and that of other major emerging economies in Latin America. This is true at all levels, but in a higher proportion in case of upper secondary and tertiary education (Figure 3). Moreover, a significant amount of public funding is absorbed by teachers' salaries (83%, compared to the OECD average of 63%). This leaves a limited fund for investment in resources that could improve quality, such as teacher's training and text books. There are also scarce resources for the renovation and expansion of education infrastructure: 1.7% of spending goes to capital expenditure compared with 7% in the OECD (OECD, 2014a), or for spending on food, transportation and accommodation, which would be also important for improving education quality in Colombia.

Figure 3. **Annual expenditure per student by educational level (2011)**



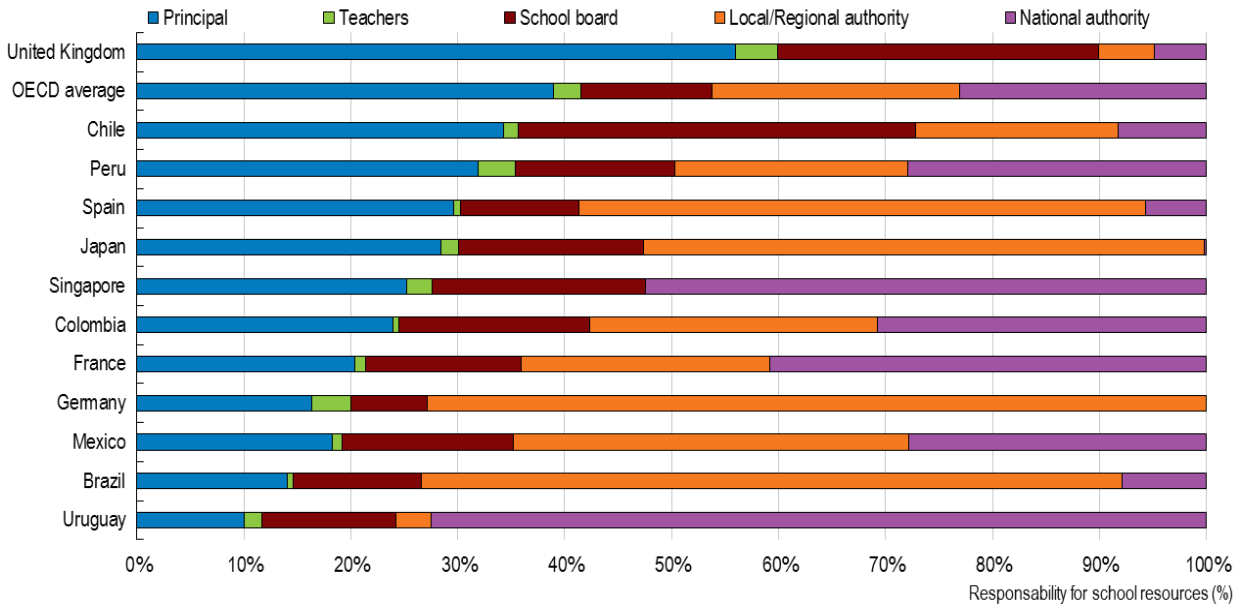
1. Year of reference 2012.

2. Public institutions only (for Colombia in tertiary education only).

Source: OECD (2014).

21. In Colombia, schools have less autonomy than the average school of OECD countries. For instance, Colombian principals and teachers hold 24% of the responsibility for resources, compared to 42% of that responsibility across OECD countries; 61% of the responsibility for the curriculum, compared to 66% across OECD countries; and 31% of the responsibility for student assessment, compared to 68% across OECD countries (Figure 4) (OECD, 2016d).

Figure 4. **Distribution across the education system of responsibility for school resources**



Note: Countries and economies are ranked in descending order of the responsibility held by school principals and teachers.

Source: OECD (2016d).

Data

22. The quantitative analysis developed in this paper relies on score data of the Saber 11 test conducted by the Colombian Institute for the Promotion of Higher Education (ICFES). This test is required for admission to higher education in Colombia, it takes place twice in a year and it is the oldest national examination in the country; it was first implemented in 1968.

23. We use in this paper data from 2015. It is important to highlight this since the ICFES made some changes on Saber 11 in 2014 with the purpose of making comparable the Saber tests for different schooling grades. For our target year, the test is divided in two sessions of 4 hours and 30 minutes each of them. It comprises a series of exams that measure 11th grade students' performance in the following areas: Spanish language, Mathematics, Natural Sciences (physics, chemistry and biology), Social Sciences and English language. The grade in every area qualifies from 0 to 100, and the total grade from 0 to 500. Moreover, the ICFES develops a global index in order to establish a grade classification and classifies students in 1,000 spots according to their score. The index is calculated as follows.

$$GLOBAL\ INDEX = \frac{3 * \mathrm{SEL} + \ 3 * \mathrm{SM} + 3 * \mathrm{SSS} + 3 * \mathrm{SNS} + 1 * \mathrm{SE}}{13}$$

24. where SEL is the score in Spanish Language, SM is the score in Mathematics, SSS the score in Social Sciences, SNS the score in Natural Sciences and SE is the score in English.

25. Despite the existence of a classification of grades, there is not a minimum one to pass the exam. Universities are free to set their own criteria about what is the minimum grade for being accepted as tertiary education student. However, general consensus exists about how to interpret the grade in terms of academic performance: grades between 0 and 30.99 are considered low; those between 31 and 45.99 are low-medium, between 46 and 70.99 high-medium and those higher than 71 are considered high.

26. We combine Saber 11 test results with some data extracted from the Ministry of National Education (MEN) as well as the National Department of Statistics (DANE) in order to develop a broad analysis.

27. After gathering this data, we develop a descriptive and quantitative analysis, using econometric tools. We estimate linear regressions following the ordinary least squares (OLS) method with the purpose of investigating the determinants of academic performance in Saber 11 test in Colombia.

28. As a dependent variable, we take the Saber 11 score in Mathematics and Language, measured on a scale from 0 to 100. We focus on these two areas since a good competence in any of them are key to succeeding professionally in any work area, while the other areas covered in Saber 11 are more specific and useful in concrete work environments. Moreover, these areas are also covered by the Programme for International Student Assessment (PISA), a cross-country test developed by the OECD in order to evaluate students' performance on Mathematics, Science and Language and open the possibility for international comparisions.

29. Our independent or explanatory variables cover determinants of academic performance from three different levels: individual or socioeconomic, school and institutional level. In the individual level, we analyse student's characteristics (sex, age and house location) as well socioeconomic factors such as: mother's education, monthly household income and number of books in the home. We also use two variables as controls: father's occupation and size of the household, measured by the number of people living with the student.

30. At school level, we research the influence of: the school day, full-time or other schedule; the type of school, public or private[2]; and how costly is the school, considering tuition fees.

31. The institutional sphere is covered here at department level with a focus on two variables: teacher's education, measured through the percentage of teachers holding postgraduate diploma, and public spending on education per student[3]. Apart from covering the thirty-two departments of Colombia, we also include separate data for the following Certified Territorial Entities (ETCs) and high-populated cities: Barranquilla, Bogotá, Cali, Medellín and Manizales[4]. Therefore, we account for data varying across thirty-seven Colombian departments and selected ETCs.

32. After introducing the variables, it is important to show the main features of our data in order to enable the reader to interpret the results of the empirical part in the right way. First of all, it is worth mentioning we only keep in the Saber11 test database those individuals born from 1996 to 1998. This is students who have between sixteen and eighteen years old at the time they took the test. Doing this we aim to isolate correctly the effect of the already mentioned explanatory variables on academic performance, reducing the probability of assigning wrongly values to explanatory variables. For instance, a thirty-year-old individual who took the Saber11 exam in 2015 could be influenced by factors explaining its academic performance which are not included in our analysis.

33. Taking into account the previous consideration, we show in Figure 5 the average score in Mathematics and Language by department. At national level, the average grade is very similar in both subject, being 50.64 for Mathematics and 50.27 for Language. The average difference between subjects across departments is also small, 0.16 grade points.

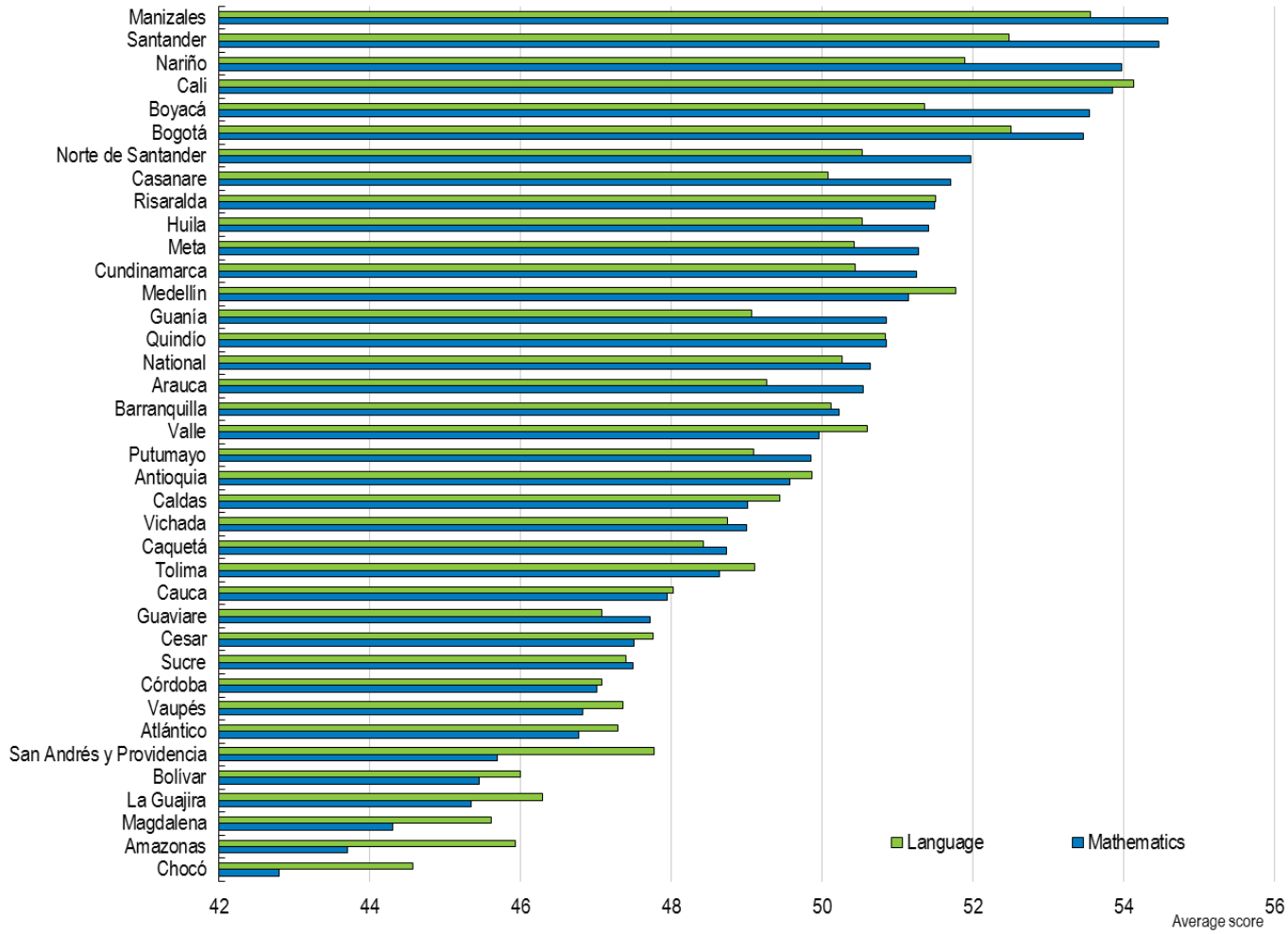
34. However, we find enormous differences between departments when we compare grades on Mathematics and Language separately. The score distance between the department accounting for the lowest grade in Mathematics (Chocó) and the one holding the highest (Manizales) is 11.78 grade points. This distance is lower but significant also in Language, reaching 9.54 grade points. This shows the huge variability on student's academic performance across departments in Colombia.

[2] At this stage, it is important to clarify that we consider public schools only those institutions entirely financed and managed by the government whereas we include under the private schools category those institutions financed and managed by private agents as well as those schools financed by the government but managed by private individuals such as *escuelas concesionadas* or *matriculas contratadas*. The election of this classification responds to the objective of analysing the effect of public spending on education on academic performance using a comparable set of schools, in this case fully public in both financial and management areas. Public schools and privated managed schools follow different dynamics, being students' average scores of the latest higher (50 for students of the former and 54 for the latest in the case of both Language and Mathematics), and more comparable with the ones obtained by private institution students (57 in Language and 59 in the case of Mathematics). For this reason, we include both *escuelas concesionadas* and *matriculas contratadas* schools in the category private.

[3] This is total public spending per student in education, including funds from the Central and regional governments forall education levels.

[4] We only include data for the mentioned ETCs as most upper-secondary education institutions are concentrated in the capital of each department, making a departamental analysis more addequate rather than an ETC approach.

Figure 5. **Average score in Mathematics and Language by department**



Source: Saber 11 test, 2015 results (ICFES).

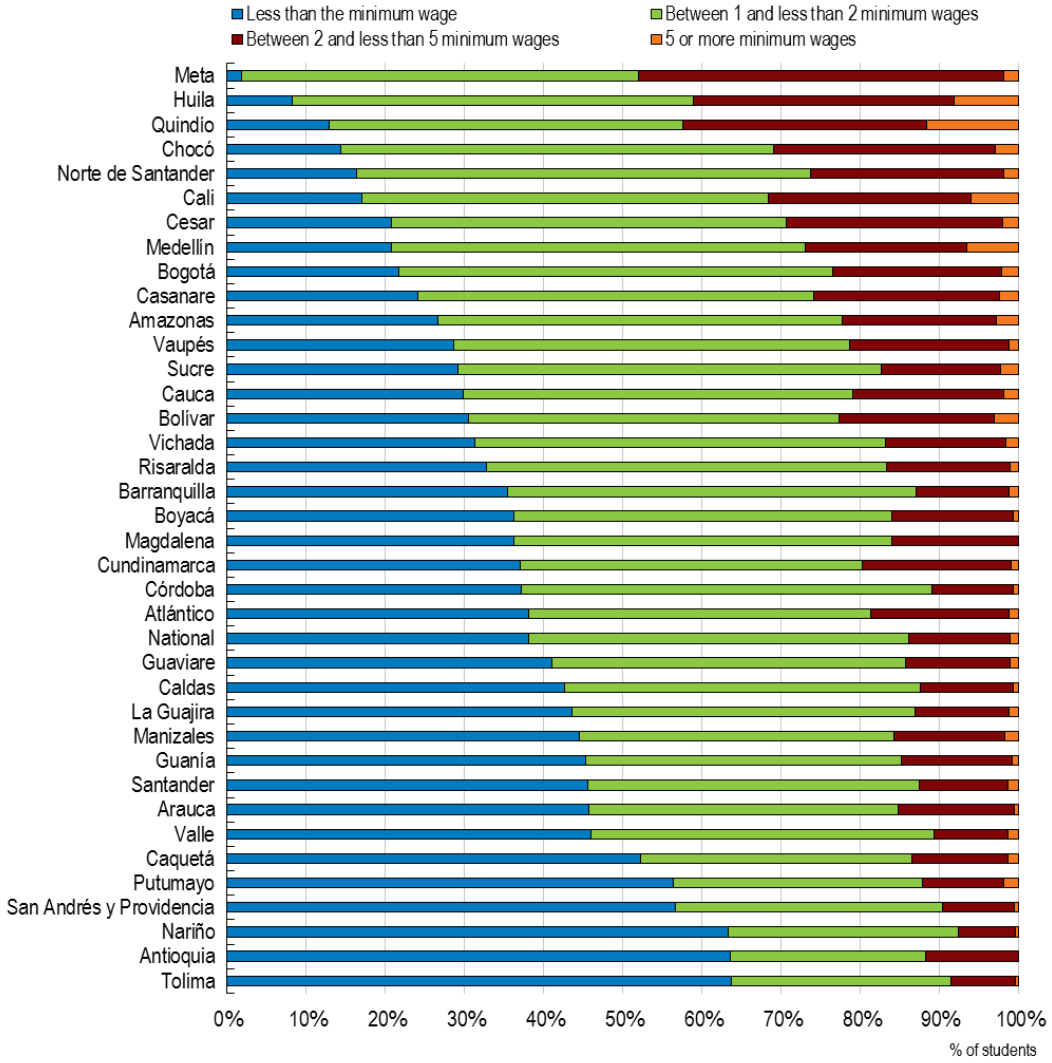
35. Mathematics and Language grades also show differences when we consider sex. Boys perform 3.4 grade points higher than girls in Mathematics (52.54 grade points the former and 49.14 the latter) and have also a slight advantage over girls in Language, scoring 0.1 grade points higher (50.34 grade points for boys and 50.23 for girls). This looks like a puzzling finding since the literature on education highlights girls usually perform better than men in Language, while the latter ones account for an advantageous position in Mathematics. This seems not to be the case of Colombia or at least not the case of Saber11 test in 2015, where girls performed worse than boys both in Mathematics and Language. This is even more significant if we take into consideration the fact that girls are the 55.42% of Saber11 test takers.

36. One feature we would like to stress is the distribution of Saber11 test takers by their household's income. We show it in the Figure 6. We aggregate some household income groups in order to make the graph more comprehensive. As we can see, 86.11% of individuals belong to households earning less the minimum wage in Colombia[5] or households earning between one and less than two minimum wages, at

[5] The minimum wage was 644,350 Colombian peso in 2015 (Data extracted from Central Bank of Colombia: http://obiee.banrep.gov.co/analytics/saw.dll?Go&Path=/shared/Consulta%20Series%20Estadisticas%20desde%20Excel/1.%20Salarios/1.1%20Salario%20minimo%20legal%20en%20Colombia/1.1.1%20Serie%20historica&Options=rdf&NQUser=salarios&NQPassword=salarios&lang=es0).

national level. For its part, those individuals belonging to wealthy households, this is those households that earn an income of five or more minimum wages, account for only the 1.06% at national level and are concentrated mainly in big cities. This is related to the fact that the majority of test takers live in Bogotá (50.39%) and the percentage reaches 76.34% when considering the four most populated cities in Colombia[6].

Figure 6. **Distribution of test takers by household income**

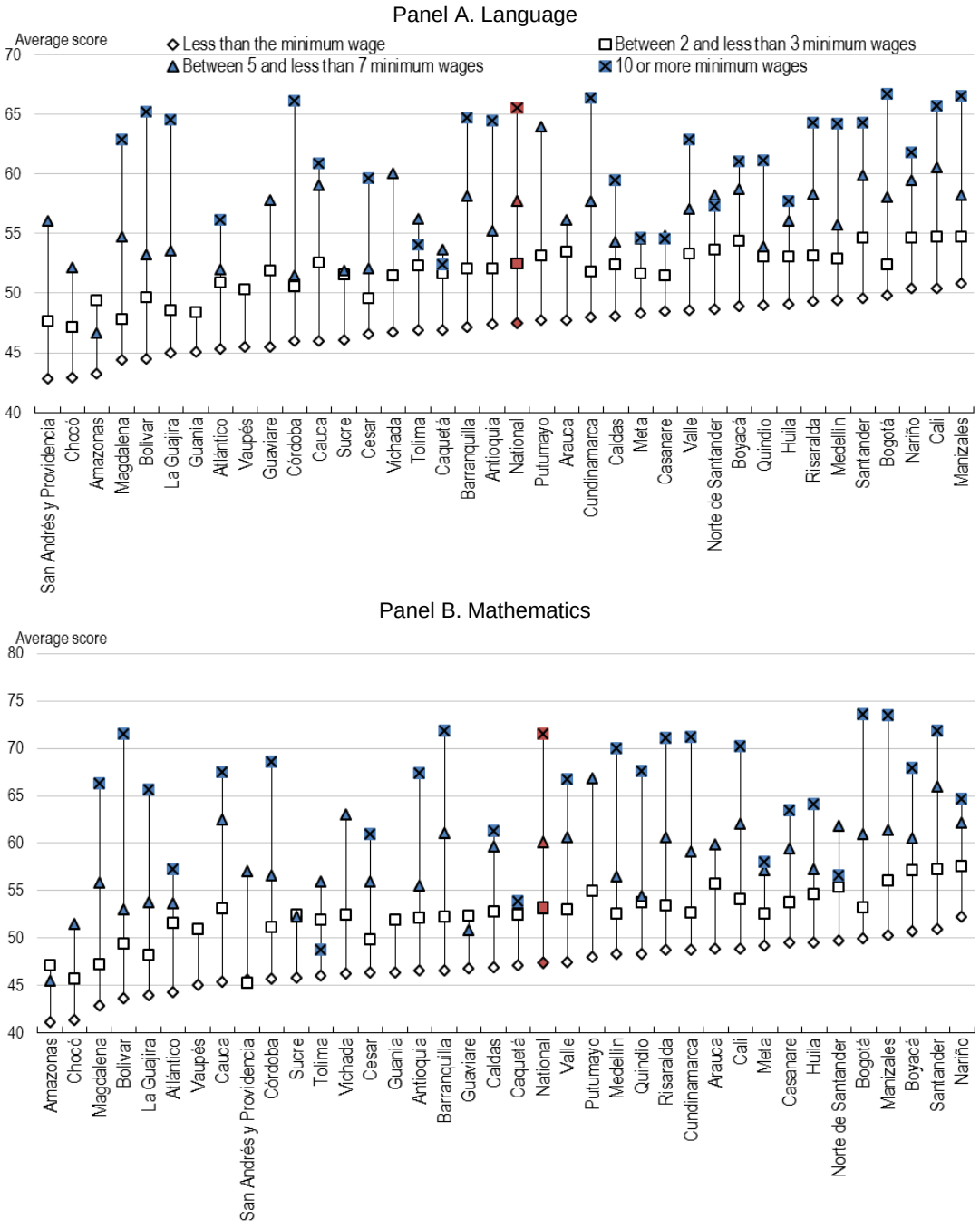


Source: Saber 11 test, 2015 results (ICFES).

37. We also can take a look at grade differences between income groups. In the figure 7 we show the average score for Language and Mathematics by household income and department. We selected four income groups, two from the extremes and two from the centre of the distribution. As we can observe, there are important differences in terms of academic performance in both subjects between them. At national level, this difference reachs more than eighteen grade points for Language and twenty-four for Mathematics.

[6] This is Bogotá, Barranquilla, Medellín and Cali.

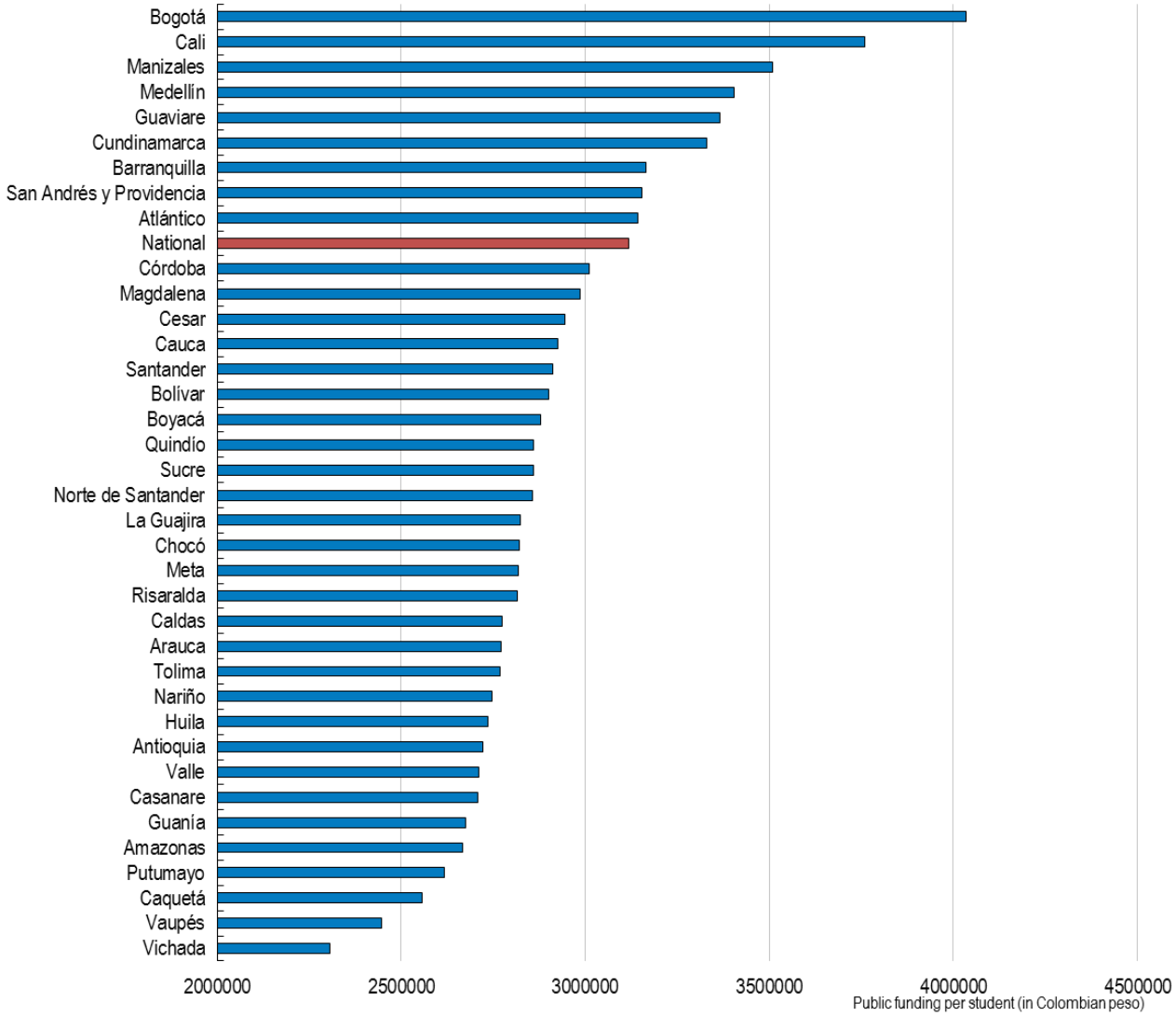
Figure 7. **Average score by monthly household income and department**



Source: Saber 11 test, 2015 results (ICFES).

38. However, we find stronger variations at the department level. When we consider only departments accounting for individuals in all the income groups represented here, we observe the smallest variations between the extreme income groups in the departments of Tolima (2.70 grade points) and Caquetá (6.73 grade points). The highest ones appear in the department of Bolívar (27.89) and the city of Barranquilla (25.24). In general, all big cities included in the sample show a great dispersion between those individuals coming from a low economic background and those accounting for a high economic background. On the other side, those students coming from low-income departments show less dispersion between income groups and in some cases there is a lack of individuals from the highest income group. Therefore, the homogeneity of individuals in those regions explain partly the lower score dispersion across income groups.

Figure 8. **Public funding per student by department**



Source: OECD (2016a).

39. Since the purpose of this research paper is addressing the relation between public spending in education and academic performance it is a key to show the distribution of the former across Colombian departments. At national level, Colombia invested 3,117,233 pesos by student in 2014 (OECD, 2016a).

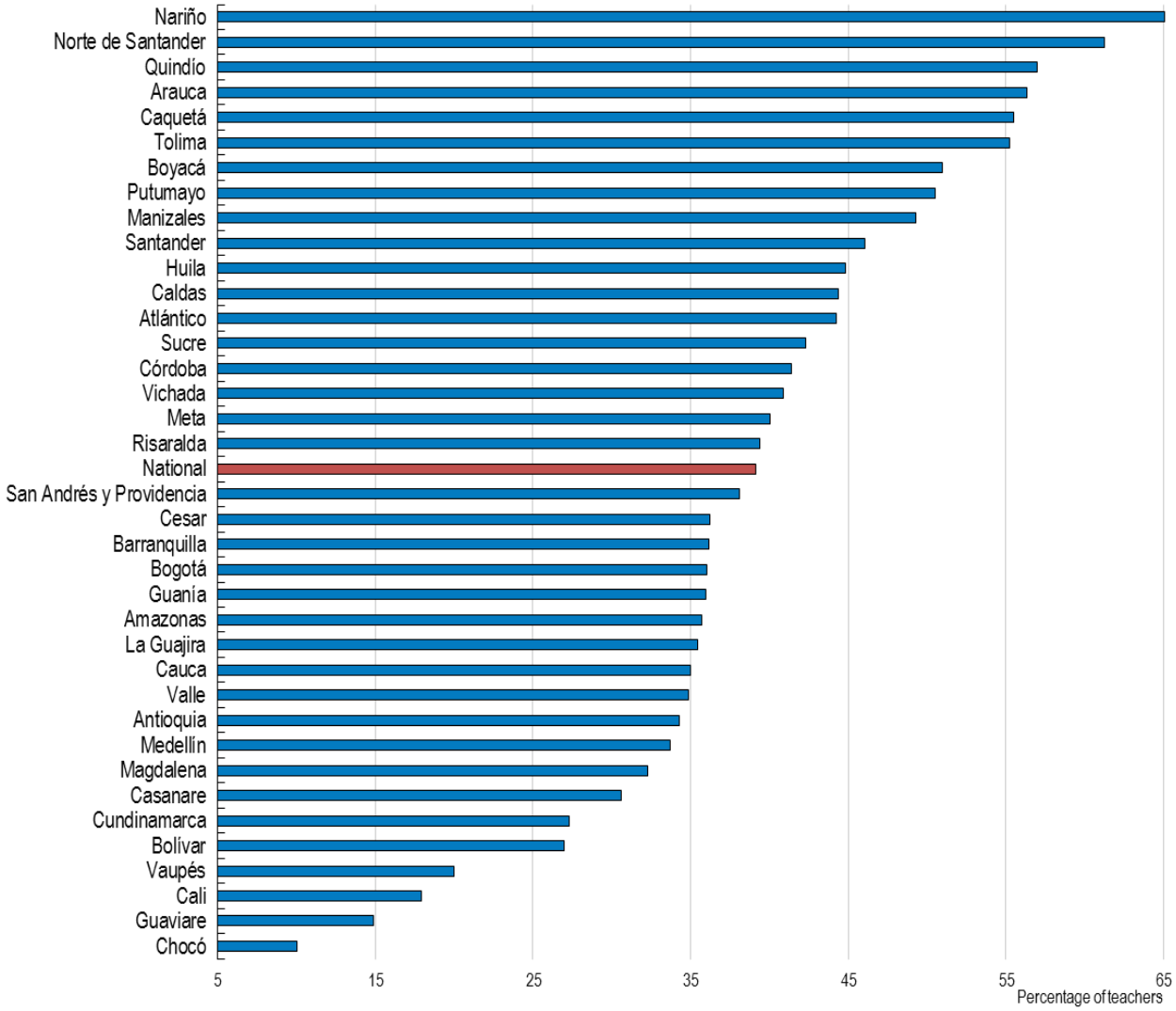
However, we find significant gaps between departments (Figure 8). The spending in education is led by the capital of Colombia, Bogotá, and followed by some of the biggest cities of the country (Cali, Manizales and Medellín). On the other side, the departments who spent the least are those located in the Amazon region (Vichada, Vaupés and Caquetá), a low populated and rural area of Colombia. For instance, Bogotá invests in education 1.7 times more than Vichada, which gives us an idea about the important inequalities Colombian departments face regarding public funding on education.

40. Teaching quality is also an important factor for ensuring a quality education and therefore a successful academic performance of students. Teacher qualifications are analysed here as proxy of it although empirical evidence increasingly shows that both qualification and skills matter for ensuring good education (OECD, 2013b). As we already mentioned, poor teachers' qualifications hinder students' academic learning and in the end compromise the economic development of a country (Bruns & Luque, 2015). We measure teachers' qualifications using the percentage of teachers holding a postgraduate diploma in this paper. This selection responds to the fact that, according to the data provided by the DANE, 97.5% of teachers of secondary education[7] reached at least a graduate level of education in Colombia. Therefore, filling the gap for teachers' qualification in Colombia involves increasing the number of teachers pursuing postgraduate studies. This is why, the percentage of teachers holding postgraduate diploma is a common variable in the existing literature about academic performance determinants in Colombia (Gaviria and Barrientos, 2001; Bonilla and Galvis, 2014).

41. When looking at proportion of teachers holding postgraduate diploma in Colombia, we find that this is the case of 39% of teachers on national average (Figure 9). Nonetheless, we find again several disparities across Colombian departments. While in some departments postgraduate educated teachers' accounts for more than 60% of total teachers (Nariño and Norte de Santander), others do not reach 20% (Chocó, Guaviare and Cali). In general, we observe that the lowest relative share of postgraduate educated teachers are concentrated in some of the poorest departments of Colombia (Chocó, Guaviare and Vaupés).

[7] "*Educación media"* in Spanish. This is the last two years of secondary education (10^{th} and 11^{th} grades).

Figure 9. **Percentage of teachers holding a postgraduate diploma**



Source: National Department of Statistics (DANE) (https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/poblacion-escolarizada/educacion-formal).

Empirical strategy

42. The analysis proposes a model where student performance in Mathematics and Language, respectively, are the dependent variables and the socioeconomic background of student, sex, age and school characteristics are included as independent variables. The estimation also exploits differences in policy indicators across regions, and employs interactions to study if more investment in public education could help reduce the impact that socioeconomic factors, monthly household income, have on student performance. Moreover, we analyse the role of teacher's qualifications for reducing the impact that monthly household income has on student's academic performance. To tackle those questions we follow the ordinary least squares (OLS) method including interactions of our two policy variables, public spending in education per student and percentage of teachers holding a postgraduate diploma, with the socioeconomic level of students, measured through monthly household income. In the case of public spending in education we use the following the equation:

$$P_i = \beta_1 Inc_i + \beta_2 ME_i + \beta_3 Sex_i + \beta_4 Age_i + \beta_5 PSE_j + \beta_6 TE_j + \beta_7 Inc_i * PSE_j + A\theta_i + \varepsilon_i$$

43. where P_i represents the SABER 11 score in Mathematics (or Language) for student *i*, Inc_i is monthly family income, ME_i is the level of education of the student's mother, *Sexi* is a dummy variable for sex taking the value 1 for girls, Age_i is the age of student *i*, *A* is a matrix containing the identified "traditional" variables (e.g. school characteristics), PSE_j is the level of public spending per student that varies at the regional level *j*, TE_j is the percentage of teachers holding postgraduate diploma at regional level *j*, and ε_i is an error term.

44. In the case of the interaction between teacher's qualifications and student's academic performance we use the following equation:

$$P_i = \beta_1 Inc_i + \beta_2 ME_i + \beta_3 Sex_i + \beta_4 Age_i + \beta_5 PSE_j + \beta_6 TE_j + \beta_7 Inc_i * TE_j + A\theta_i + \varepsilon_i$$

45. We complement this analysis using a probit model in order to focus on the factors behind the low-performing in Mathematics and Language at Saber 11 test. For this purpose, we consider low-performers to those students scoring below 50 in Mathematics or Language and use for the estimation the following equation:

$$P_i(Y = 1 \mid X) = \beta_1 Inc_i + \beta_2 ME_i + \beta_3 Sex_i + \beta_4 Age_i + \beta_5 PSE_j + \beta_6 TE_j + A\theta_i + \varepsilon_i$$

46. Where P_i is a binary variable that takes the value of 1 when the student *i* scored below 50 score points in the subject (Mathematics and Language respectively) or the value of 0 when the test taker scored 50 or higher, Inc_i is monthly family income, ME_i is the level of education of the student's mother, *Sexi* is a dummy variable for sex taking the value 1 for girls, Age_i is the age of student *i*, *A* is a matrix containing the identified "traditional" variables (e.g. school characteristics), PSE_j is the level of public spending per student that varies at the regional level *j*, TE_j is the percentage of teachers holding postgraduate diploma at regional level *j*, and ε_i is an error term.

Results: the determinants of student's performance

47. In this section, we present the results of our quantitative analysis in order to explain academic performance in the Saber 11 test and explore the potential benefit of increasing public spending on education and teacher qualifications for reducing the effect of the socioeconomic status of students on academic results. We are also interested in the role that public spending on education could play for reducing the socioeconomic differences among students and for ensuring that all can benefit from education equally.

48. When looking at the results of the baseline estimation for Mathematics and Language (Table 1) we can get a complete insight into the individual, school and departmental determinants of academic performance in Saber 11 for the year 2015.

Table 1. **Baseline estimations**

Panel A. Mathematics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent variable: Mathematics score											
Female	-3.401***	-3.399***	-3.579***	-3.335***	-3.162***	-3.187***	-3.206***	-3.207***	-3.267***	-3.268***	-3.245***
	(0.0374)	(0.0374)	(0.0380)	(0.0365)	(0.0357)	(0.0354)	(0.0352)	(0.0352)	(0.0351)	(0.0351)	(0.0348)
Age		0.136***	0.145***	0.0941***	-0.00818	-0.00610	-0.0214	-0.0202	0.0110	0.0120	-0.0233
		(0.0243)	(0.0249)	(0.0239)	(0.0234)	(0.0232)	(0.0231)	(0.0231)	(0.0230)	(0.0230)	(0.0229)
Urban area			3.676***	2.079***	1.551***	1.371***	1.768***	1.766***	2.038***	2.007***	1.899***
			(0.0512)	(0.0500)	(0.0491)	(0.0487)	(0.0491)	(0.0491)	(0.0491)	(0.0495)	(0.0492)
Mother's education				1.603***	0.960***	0.804***	0.793***	0.794***	0.758***	0.761***	0.781***
				(0.00929)	(0.0105)	(0.0106)	(0.0105)	(0.0106)	(0.0105)	(0.0106)	(0.0105)
Monthly household income					-2.240***	-1.755***	-1.627***	-1.635***	-1.253***	-1.237***	-1.255***
					(0.0198)	(0.0205)	(0.0205)	(0.0217)	(0.0226)	(0.0229)	(0.0227)
Access to books in the home						1.843***	1.782***	1.781***	1.753***	1.741***	1.667***
						(0.0223)	(0.0222)	(0.0223)	(0.0221)	(0.0223)	(0.0222)
Full-time schooling							2.793***	2.802***	2.132***	2.155***	2.061***
							(0.0527)	(0.0534)	(0.0544)	(0.0546)	(0.0543)
Public status of the school								0.0928	7.955***	7.985***	7.712***
								(0.0892)	(0.162)	(0.162)	(0.161)
School tuition fees									2.742***	2.737***	2.791***
									(0.0472)	(0.0472)	(0.0469)
log(Public spending per student)										0.658***	4.059***
										(0.136)	(0.144)
% of teachers holding a postgraduate diploma											0.116***
											(0.00175)
Constant	52.54***	50.26***	46.87***	41.84***	62.61***	57.29***	56.13***	56.07***	42.65***	32.70***	-21.66***
	(0.0278)	(0.409)	(0.421)	(0.405)	(0.431)	(0.432)	(0.431)	(0.435)	(0.490)	(2.114)	(2.254)
Observations	371,745	371,745	340,731	340,355	340,355	340,355	340,355	340,355	340,293	340,293	340,293
R-squared	0.022	0.022	0.039	0.116	0.155	0.171	0.178	0.178	0.186	0.186	0.196

Panel B. Language

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent variable: Language score											
Female	-0.103***	-0.100***	-0.283***	-0.0777***	0.0662**	0.0473*	0.0357	0.0386	-0.00501	-0.00738	0.00449
	(0.0295)	(0.0295)	(0.0303)	(0.0290)	(0.0284)	(0.0281)	(0.0280)	(0.0280)	(0.0280)	(0.0279)	(0.0279)
Age		0.189***	0.0801***	0.0381**	-0.0443**	-0.0427**	-0.0520***	-0.0593***	-0.0369**	-0.0346*	-0.0527***
		(0.0192)	(0.0199)	(0.0190)	(0.0186)	(0.0184)	(0.0184)	(0.0184)	(0.0183)	(0.0183)	(0.0183)
Urban area			3.216***	1.885***	1.452***	1.313***	1.556***	1.566***	1.760***	1.692***	1.637***
			(0.0409)	(0.0398)	(0.0390)	(0.0387)	(0.0391)	(0.0391)	(0.0391)	(0.0394)	(0.0394)
Mother's education				1.337***	0.814***	0.694***	0.687***	0.681***	0.655***	0.662***	0.673***
				(0.00739)	(0.00833)	(0.00839)	(0.00837)	(0.00841)	(0.00840)	(0.00841)	(0.00839)
Monthly household income					-1.774***	-1.402***	-1.324***	-1.281***	-1.009***	-0.973***	-0.982***
					(0.0157)	(0.0163)	(0.0163)	(0.0173)	(0.0180)	(0.0182)	(0.0182)
Access to books in the home						1.415***	1.378***	1.379***	1.359***	1.333***	1.295***
						(0.0177)	(0.0177)	(0.0177)	(0.0177)	(0.0178)	(0.0177)
Full-time schooling							1.709***	1.657***	1.179***	1.229***	1.180***
							(0.0420)	(0.0426)	(0.0434)	(0.0435)	(0.0434)
Public status of the school								-0.530***	5.077***	5.141***	5.002***
								(0.0710)	(0.129)	(0.129)	(0.129)
School tuition fees									1.955***	1.945***	1.973***
									(0.0376)	(0.0376)	(0.0375)
log(Public spending per student)										1.425***	3.167***
										(0.108)	(0.116)
% of teachers holding a postgraduate diploma											0.0594***
											(0.00140)
Constant	50.34***	47.16***	46.10***	41.90***	58.98***	54.89***	54.18***	54.54***	44.97***	23.41***	-4.423**
	(0.0220)	(0.323)	(0.336)	(0.322)	(0.343)	(0.343)	(0.343)	(0.346)	(0.391)	(1.685)	(1.803)
Observations	371,745	371,745	340,731	340,355	340,355	340,355	340,355	340,355	340,293	340,293	340,293
R-squared	0.000	0.000	0.018	0.104	0.145	0.161	0.165	0.165	0.171	0.172	0.176

Note: Asterisks (*, **, ***) indicate de significance level (10%, 5% 1%) of the coefficients.

49. In our model, the individual level contributes to explain 17% of variation in academic performance across students in Saber 11, being the most relevant level for explaining the phenomenon. When analysing the variables covered at this level, we can highlight boys seem to perform better than girls in the two main areas covered by Saber 11, Mathematics and Language. In Mathematics the difference between sexes reaches as much as three score points while in the case of Language, boys hold a small advantage over girls although this variable turns non-significant when controlling for more variables in the model. Thus, our finding contradict the literature on the topic since our model does not show an advantage for girls in Language.

50. Regarding the age of test takers, we do not find a significant link with academic performance, but it is important to note that we are considering for the estimation only those individuals who were between 16 and 18 years old at the time they took the test. Students living in rural areas face a disadvantaged position compared to those living in cities. Particularly, those living in cities score 1.90 points higher in Mathematics than those living in rural areas and 1.64 points in Language, when controlling by all aspects covered in the model.

51. Socioeconomic characteristics of the student also play an important role for assessing academic performance differences across individuals and are addressed here using three variables: mother's education, monthly household income and access to books in the home. For the first one, increasing mother's education by one, this is moving from one category to the next higher one, is associated with a 0.78 points higher score in Mathematics and 0.67 in Language when we control for all variables.

52. Monthly household income is also a categorical variable that classifies individuals in seven groups following a descending order. This is, category 7 is assigned to those individuals living in households reporting an income lower than the minimum wage and in the other extreme, category 1 covers those individuals belonging to households where the monthly income is higher than ten minimum wages. This variable shows therefore a negative coefficient in the models which mean transiting from one category to the immediate higher one reduce the score 1.26 points in Mathematics and 0.98 in Language in the case we include all aspects influencing academic performance.

53. The access to books in the home is also associated with better academic outcomes. In this case, jumping from one category to the following one, this is increasing the number of books the student have access at home, is associated with a 1.67 points higher score in Mathematics and 1.30 in Language when controlling for all variables.

54. With regards to school characteristics, this level explains a little bit more than 1% of disparities across individuals in terms of academic performance. Focusing on the variables, we can point out full-time schooling could help increase student's score by 2.06 points in Mathematics and 1.18 in Language. The public status of the school does not seem to be significant in the model for Mathematics. The estimated coefficient has a negative sign for Language but when controlling for school tuition fees it turns positive and significant in both areas. Specifically, students going to public schools score 7.71 points higher in Mathematics than those going to private ones, and 5.00 points in Language. This finding is in line with the existing literature about private schooling in Colombia (García et al., 2013 & OECD, 2016b), which shows that going to a private school in itself does not guarantee achieving good academic results but going to a high-fees private school does.

55. Departmental variables, public spending per student and percentage of teachers holding postgraduate diploma can explain 1% of variation across test takers of Saber 11. Public spending on education, introduced in the model with a log, can play an important role for improving academic performance in Colombia. When controlling for the rest of variables, increasing public spending per student by 1% is associated with a 4.05 points higher average score in Mathematics and 3.17 points in

Language. Increasing the percentage of teachers holding postgraduate diploma is also positively related to academic scores, although the estimated coefficient is smaller. In this case, increasing the proportion of teachers holding postgraduate diploma by 1 would be consistent with a 0.12 points higher average score in Mathematics and 0.06 points in Language.

56. After this brief overview of variables related to academic performance in Saber 11 we will focus on the role of public spending in education for reducing the effect of household income on academic outcomes. The choice of this variable responds to the fact that monthly household income can explain the highest part of academic differences across individuals (4%).

Interaction terms: measuring the role of public spending on education and teacher qualifications

57. Increasing both public funding in education and teacher qualifications can be a good way to improve academic performance, as suggested by the analysis presented above, but apart from that it also could help to reduce the role of socioeconomic characteristics of students such as their household income. We include interaction terms in the baseline models in order to explore this issue.

58. Before presenting the results we need to make clear the following considerations. First, we include in the model only those individuals attending to public schools since private schools do not receive public funding and therefore their students are not influenced by changes in public funding in education. Second, after restricting the model only to those students attending public schools we capture significant changes in the composition of the variable monthly household income. In particular, categories 1, 2 and 3 lose 8,656 individuals in total, decreasing from 11,362 individuals when including private school students to 2,706 individuals when considering only those attending public institutions. For this reason we aggregate the mentioned categories in only one that represent those students belonging to families whose monthly income is 5 or more minimum wages. The rest of categories for this variable remain the same. Therefore, monthly household income is here a variable composed of five categories in the following models: 1 for referring to those individuals whose household income is higher than 5 minimum wages, 2 for those between 3 and less than 5 minimum wages, 3 for those between 2 and less than 3 minimum wages, 4 for those between 1 and less than 2 minimum wages and 5 for those whose income is less than the minimum wage.

59. The use of interactions between monthly household income and public spending per student as well as between monthly household income and percentage of teachers holding postgraduate diploma allows us to analyse the role of these policy variables from two perspectives[8]. On the one hand, we can interpret the interaction from the point of view of the potential role of public spending per student for reducing the impact of monthly household income on academic performance, i.e. to what extent public spending in education could help diminishing the differences in student academic results arising from different socioeconomic backgrounds. On the other hand, we can understand the interaction from the angle of exploring to what extent public spending on education can have a differential impact on academic performance depending on the socioeconomic background of the student.

[8] In order to make it clear we explain it here the case of public spending per student.

Table 2. **Baseline estimation with interactions for Mathematics**

	(1)	(2)	(3)
Dependent variable: Mathematics score			
Female	-3.256***	-3.257***	-3.257***
	(0.0355)	(0.0355)	(0.0355)
Age	0.0190	0.0192	0.0195
	(0.0233)	(0.0233)	(0.0233)
Urban area	2.057***	2.054***	2.063***
	(0.0499)	(0.0500)	(0.0499)
Mother's education	0.778***	0.778***	0.781***
	(0.0107)	(0.0107)	(0.0107)
Monthly household income	-1.153***	-1.884***	-1.675***
	(0.0248)	(0.272)	(0.0809)
Access to books in the home	1.686***	1.686***	1.687***
	(0.0227)	(0.0227)	(0.0227)
Full-time schooling	1.823***	1.824***	1.820***
	(0.0559)	(0.0559)	(0.0559)
log(Public spending per student)	3.891***	3.605***	3.815***
	(0.148)	(0.182)	(0.149)
% of teachers holding a postgraduate diploma	0.120***	0.120***	0.0381***
	(0.00178)	(0.00178)	(0.0122)
log(Public spending per student)*Monthly household income		0.0482***	
		(0.0179)	
% of teachers holding a postgraduate diploma*Monthly household income			0.0132***
			(0.00195)
Constant	-12.50***	-9.635***	-9.212***
	(2.290)	(2.525)	(2.340)
Observations	321,946	321,946	321,946
R-squared	0.136	0.136	0.136

Note: Asterisks (*, **, ***) indicate de significance level (10%, 5% 1%) of the coefficients.

60. From the first perspective, the aggregate effect of monthly household income can be interpreted as follows. This is the coefficient of the variable covering the economic status of students plus the coefficient of the interaction term multiplied by the different values of the policy variable, log(public spending per student). In the case of Mathematics (Table 2), when the latest variable reaches its higher value (15.21) the aggregate effect of monthly household income is -1.178 while it is -1.151 for the lowest value of log(public spending per student). The highest value represents the department that invested the most in education, Bogotá, and is slightly lower than the coefficient of monthly household income when not introducing interaction (-1.153). Thus, it supports the idea that a high level of public spending in education could have a downward effect in score differences derived from different socioeconomic background of test takers. Nonetheless, the difference is very small and could suggest that the level of public spending in education in Colombia is too low to correct the differences arising from the socioeconomic background of students and it should be better accompanied by mechanisms to strengthen its effectiveness. In the case of Language (Table 3), we find the same pattern of results, with the aggregate effect of monthly household income slightly higher than in the previous case when not including the interaction.

Table 3. **Baseline estimation with interactions for Language**

	(1)	(2)	(3)
Dependent variable: Language score			
Female	-0.0626**	-0.0629**	-0.0637**
	(0.0285)	(0.0285)	(0.0285)
Age	-0.00905	-0.00892	-0.00860
	(0.0187)	(0.0187)	(0.0187)
Urban area	1.743***	1.740***	1.749***
	(0.0401)	(0.0401)	(0.0401)
Mother's education	0.664***	0.664***	0.667***
	(0.00861)	(0.00861)	(0.00862)
Monthly household income	-0.984***	-1.670***	-1.514***
	(0.0199)	(0.218)	(0.0649)
Access to books in the home	1.297***	1.297***	1.298***
	(0.0182)	(0.0182)	(0.0182)
Full-time schooling	1.026***	1.027***	1.023***
	(0.0448)	(0.0448)	(0.0448)
log(Public spending per student)	3.036***	2.768***	2.959***
	(0.119)	(0.146)	(0.119)
% of teachers holding a postgraduate diploma	0.0638***	0.0639***	-0.0194**
	(0.00143)	(0.00143)	(0.00981)
log(Public spending per student)*Monthly household income		0.0453***	
		(0.0143)	
% of teachers holding a postgraduate diploma*Monthly household income			0.0134***
			(0.00156)
Constant	1.713	4.406**	5.051***
	(1.837)	(2.026)	(1.878)
Observations	321,946	321,946	321,946
R-squared	0.116	0.116	0.116

Note: Asterisks (*, **, ***) indicate de significance level (10%, 5% 1%) of the coefficients.

61. We follow the same procedure to analyze the aggregate effect when interacting monthly household income with percentage of teachers holding postgraduate diploma. In this case, the aggregate coefficient estimate varies widely depending on the value of teacher qualifications. For the highest value of the former variable (65.09), the estimated coefficient is -0.816, substantially lower than the coefficient of the socioeconomic variable when not introducing interaction (-1.153). In the case of the lowest value (10.04), the aggregate coefficient reaches -1.542. Increasing the proportion of postgraduate educated teachers seems to play a role for reducing those score differences derived from socioeconomic disparities of students. Moreover, we find similar results for Language.

62. Also, we can analyze the interaction terms from the point of view of the policy variable (public funding per student or percentage of teachers holding postgraduate diploma) and its contribution to improve educational outcomes, particularly of those who are of low socioeconomic backgrounds. For doing so we follow the previous procedure to quantify the aggregate effect of both policy variables. In the case of public funding per student, the aggregate coefficient estimate for Mathematics results is 3.653 when monthly household income takes value 1, this is for the richest group of students and, 3.846 when the mentioned variable takes value 5, and so considering the students that belong to families earning less than

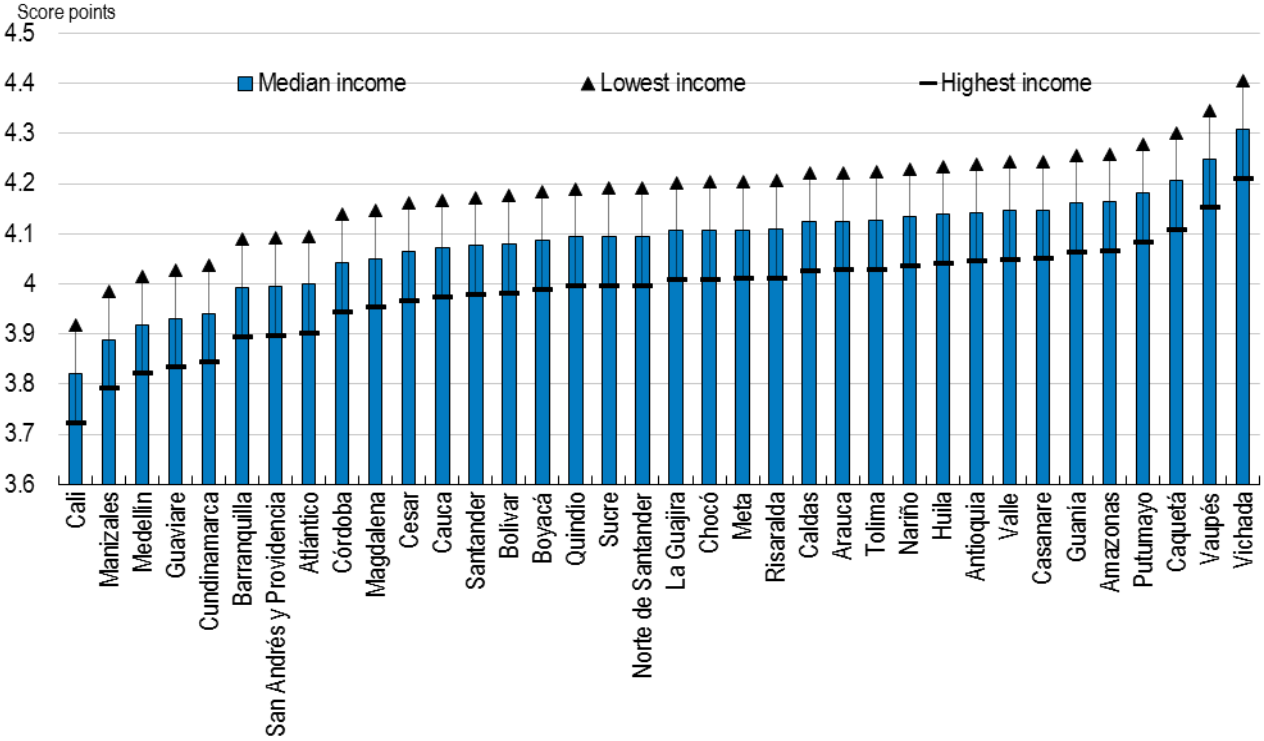
the minimum wage. The link between public spending on education and outcomes is hence not uniform and consistent with spending being particularly beneficial to those individuals from a disadvantaged socioeconomic background. We again find similar results for Language.

63. In the case of teacher qualifications, the aggregate link turns to be much higher when considering the poorest income group (0.104) than in the case of the richest one (0.051). This shows again increasing teacher qualifications could be is particularly advantageous for those students belonging to poor families as it could reduce the disadvantage they holds in terms of worse academic performance than their peers from better socioeconomic status.

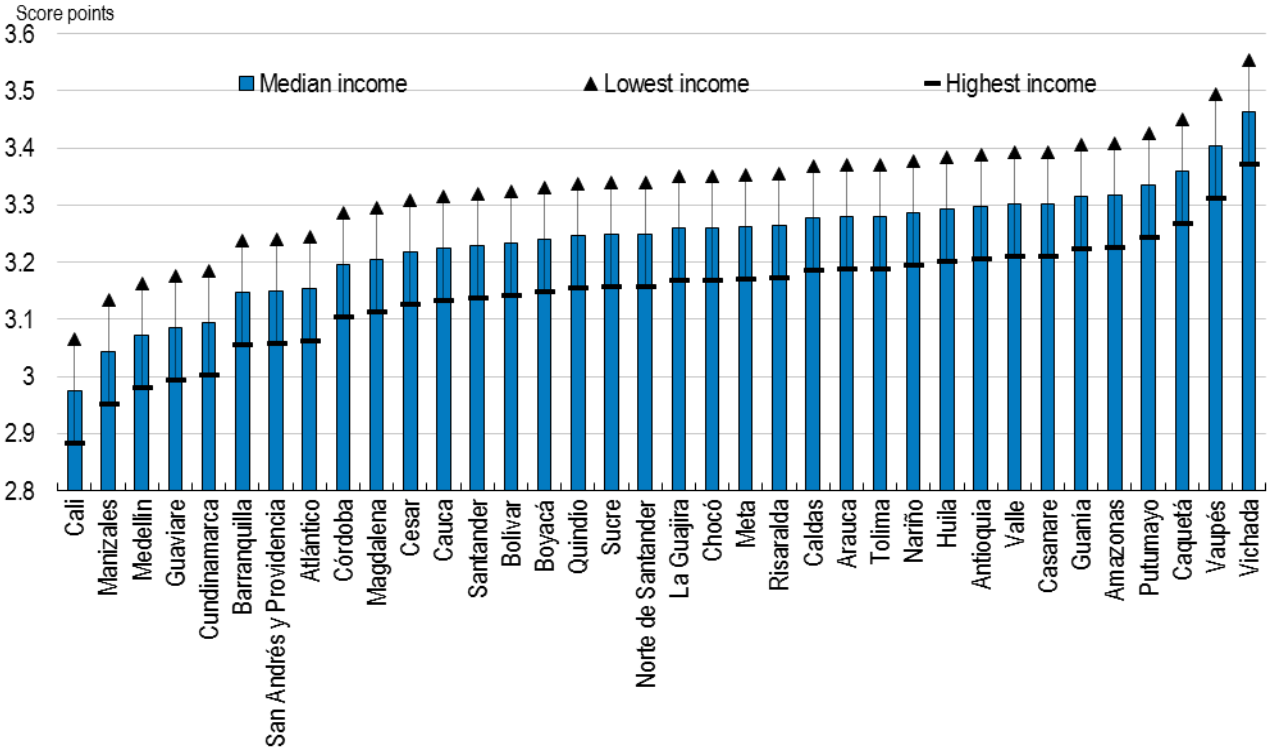
64. To deepen the measurement of the policy space the government has to improve academic performance in Colombia, we develop an empirical exercise in order to quantify the potential academic improvement in Mathematics and Language at Saber 11 test if all departments would invest at the same level than the department who do it the most, Bogotá. As we can observe at the Figure 10, doing this all departments could rise their average score in Mathematics between 3.7 and 4.4 score points and between 2.9 and 3.6 in Language. Those ones holding the worst average scores could benefit the most from this increase and particularly, low-background students could be the greatest beneficiaries.

Figure 10. **Impact of public spending on student's performance**

Panel A. Mathematics



Panel B. Language



Source: Saber 11 test, 2015 results (ICFES).

Probit estimation: determinants of low-performing in Saber 11

65. Saber 11 is a mandatory test for all senior year high school students who wish to gain access to higher education in Colombia. However, there is not a minimum score test takers need to achieve in order to pass the exam and, this fact makes difficult to set a threshold for identifying those who perform high on the test from those who show a poor performance. Moreover, we would like to emphasize that although low-performance is an important problem in all compulsory levels of education in Colombia (from preschool to 9th grade), it could be though to be less relevant in post-compulsory education or between those students willing to pursue university studies. However, after covering this issue here we can conclude this is not the case and also factors behind seem to be the same than in the case of low-performance in compulsory education.

66. Some authors consider low performers to those students scoring below 31 but here we use 50 score points as a threshold to classify students in low performers and high performers. This is not a discretional election but it responds to the fact that median score is 50.64 in Mathematics and 50.27 in Language for 2015 Saber 11 results. With this election we aim to analyze the individual, school and departmental factors behind a student's score in the inferior part of the distribution as well as the probability of being a low-performer when holding any of those characteristics.

67. We use a probit model with a dependent variable that takes value 1 when a student's score is lower than 50 or value 0 when it is 50 or higher. Estimation results are presented in Table 4 for Mathematics (Panel A) and Language (Panel B). It is worthy to remind probit coefficient cannot be interpreted except when calculating marginal effects at specific values of the other variables, since probit is not a linear model. Therefore regression coefficients can only give us an insight about the direction (negative or positive) of the independent variables impact on our explanatory variable.

Table 4. **Probit estimation results**

Panel A. Mathematics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent variable: Mathematics score											
Female	0.316***	0.316***	0.339***	0.330***	0.320***	0.326***	0.328***	0.329***	0.333***	0.334***	0.333***
	(0.00416)	(0.00416)	(0.00436)	(0.00442)	(0.00445)	(0.00447)	(0.00447)	(0.00447)	(0.00448)	(0.00448)	(0.00449)
Age		0.00168	-0.00208	0.00185	0.00761***	0.00747**	0.00808***	0.00615**	0.00309	0.00278	0.00564*
		(0.00270)	(0.00286)	(0.00290)	(0.00291)	(0.00293)	(0.00293)	(0.00293)	(0.00294)	(0.00294)	(0.00295)
Urban area			-0.381***	-0.256***	-0.208***	-0.192***	-0.221***	-0.220***	-0.236***	-0.228***	-0.217***
			(0.00596)	(0.00613)	(0.00622)	(0.00624)	(0.00633)	(0.00633)	(0.00636)	(0.00641)	(0.00643)
Mother's education				-0.136***	-0.0900***	-0.0756***	-0.0752***	-0.0769***	-0.0742***	-0.0750***	-0.0776***
				(0.00116)	(0.00132)	(0.00134)	(0.00134)	(0.00135)	(0.00135)	(0.00136)	(0.00136)
Monthly household income					0.179***	0.132***	0.127***	0.137***	0.115***	0.111***	0.115***
					(0.00264)	(0.00274)	(0.00277)	(0.00286)	(0.00296)	(0.00299)	(0.00300)
Access to books in the home						-0.183***	-0.179***	-0.179***	-0.177***	-0.174***	-0.167***
						(0.00282)	(0.00283)	(0.00283)	(0.00283)	(0.00285)	(0.00286)
Full-time schooling							-0.222***	-0.233***	-0.194***	-0.200***	-0.194***
							(0.00683)	(0.00687)	(0.00700)	(0.00703)	(0.00704)
Public status of the school								-0.156***	-0.727***	-0.735***	-0.720***
								(0.0118)	(0.0216)	(0.0217)	(0.0217)
School tuition fees									-0.219***	-0.218***	-0.229***
									(0.00700)	(0.00700)	(0.00704)
log(Public spending per student)										-0.166***	-0.499***
										(0.0174)	(0.0186)
% of teachers holding a postgraduate diploma											-0.0113***
											(0.000226)
Constant	-0.180***	-0.208***	0.181***	0.615***	-1.040***	-0.517***	-0.447***	-0.327***	0.640***	3.152***	8.484***
	(0.00310)	(0.0455)	(0.0483)	(0.0491)	(0.0539)	(0.0547)	(0.0548)	(0.0555)	(0.0633)	(0.270)	(0.291)
Observations	371,745	371,745	340,731	340,355	340,355	340,355	340,355	340,355	340,293	340,293	340,293
Pseudo R-squared	0.0112	0.0112	0.0212	0.0514	0.0643	0.0733	0.0755	0.0759	0.0781	0.0782	0.0836

Panel B. Language

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent variable: Language score											
Female	0.0166***	0.0164***	0.0385***	0.0171***	0.00114	0.00368	0.00444	0.00528	0.00852*	0.00893**	0.00730
	(0.00414)	(0.00414)	(0.00434)	(0.00441)	(0.00444)	(0.00446)	(0.00446)	(0.00446)	(0.00447)	(0.00447)	(0.00448)
Age		-0.0148***	-0.00314	0.000926	0.00682**	0.00661**	0.00701**	0.00564*	0.00250	0.00211	0.00402
		(0.00269)	(0.00284)	(0.00289)	(0.00291)	(0.00292)	(0.00292)	(0.00292)	(0.00293)	(0.00293)	(0.00293)
Urban area			-0.409***	-0.277***	-0.226***	-0.210***	-0.233***	-0.232***	-0.247***	-0.237***	-0.230***
			(0.00593)	(0.00611)	(0.00621)	(0.00624)	(0.00631)	(0.00631)	(0.00634)	(0.00640)	(0.00641)
Mother's education				-0.146***	-0.0968***	-0.0829***	-0.0826***	-0.0838***	-0.0812***	-0.0822***	-0.0838***
				(0.00117)	(0.00132)	(0.00134)	(0.00134)	(0.00135)	(0.00135)	(0.00136)	(0.00136)
Monthly household income					0.186***	0.141***	0.137***	0.144***	0.124***	0.118***	0.121***
					(0.00266)	(0.00276)	(0.00278)	(0.00286)	(0.00296)	(0.00300)	(0.00300)
Access to books in the home						-0.177***	-0.173***	-0.173***	-0.171***	-0.167***	-0.163***
						(0.00282)	(0.00282)	(0.00282)	(0.00283)	(0.00285)	(0.00285)
Full-time schooling							-0.173***	-0.180***	-0.142***	-0.150***	-0.145***
							(0.00682)	(0.00686)	(0.00699)	(0.00701)	(0.00702)
Public status of the school								-0.111***	-0.673***	-0.682***	-0.672***
								(0.0118)	(0.0215)	(0.0216)	(0.0216)
School tuition fees									-0.220***	-0.219***	-0.227***
									(0.00713)	(0.00713)	(0.00716)
log(Public spending per student)										-0.205***	-0.432***
										(0.0173)	(0.0185)
% of teachers holding a postgraduate diploma											-0.00775***
											(0.000225)
Constant	-0.0488***	0.200***	0.370***	0.842***	-0.942***	-0.434***	-0.382***	-0.295***	0.660***	3.765***	7.409***
	(0.00308)	(0.0453)	(0.0481)	(0.0489)	(0.0538)	(0.0546)	(0.0547)	(0.0554)	(0.0632)	(0.269)	(0.290)
Observations	371,746	371,746	340,731	340,355	340,355	340,355	340,355	340,355	340,293	340,293	340,293
R-squared	0.0000	0.0001	0.0103	0.0447	0.0598	0.0682	0.0696	0.698	0.0719	0.0722	0.0747

Note: Asterisks (*, **, ***) indicate de significance level (10%, 5% 1%) of the coefficients.

68. When looking at the sign of independent variables, we find similar results in both Mathematics and Language set of regressions. Individual predictors such as being a girl, living a rural area, having a low-educated mother, belonging to a low household income and lacking of access to books in the home are positively correlated with scoring below 50 in Saber 11. At the school level, attending a private, low tuition fees and part-time schooling institution correlates with low-performing. With respect to departmental characteristics, living in a department that under-invest in education and account for low qualified teachers is correlated with poor academic performance.

69. To facilitate the analysis and quantify the link between academic performance and other variables we calculate marginal effects. We show results of marginal effects at the mean in Table 4. This is the probability of an individual to be a low-performer when holding a particular feature and the rest of variables are at their average level.

70. When considering individual factors, in the case of Mathematics for instance, the predicted probability of being a low-performer when the student is a girl is 13.2% higher than for the opposite gender when the other explanatory variables hold average values. The interpretation for age would be that if age increases by 1 the probability of being low-performer increases by 0.002, which shows a small effect of this variable on the outcome. Those students living in rural areas are 8.6% more likely to be low-performers than those living in urban areas. Increasing mother's education by 1, this is moving to the immediate superior category, decreases the probability of being low-performance by 0.031 while increasing by 1 the student's monthly household income, this is moving to a lower-income category, increases the probability of being low-performer 4%. Decreasing the number of books in the home also has a positive impact on the probability of being low-performer. In particular, decreasing the number of books in the home by 1, this is moving to the immediate inferior category, raises the probability of being low-performer 6.7%.

71. At schooling level, a student attending to a full-time schooling institution has a predicted probability of being low-performer 7.7% lower than their peers enrolled in another type of institution. With respect to school status, those students attending public schools are 26.7% less likely to be low-performers than those from private institutions. Schooling tuition fees have also a negative relation with low-performing and increasing by 1 the tuition fees decreases the probability of scoring low by 0.091.

72. Departmental variables also play an important role for predicting the probability of low-performing. With regards to public funding per student, increasing its value by 1% decreases by 19.9% the probability of being low-performer. Teacher qualifications have a negative and more modest value for predicting the probability of scoring low. Particularly, increasing the percentage of teachers holding postgraduate diploma by 1 decreases the probability of being low-performer by 0.005.

Table 5. **Marginal effects**

Panel A. Mathematics

Variable	dy/dx	Std. Err.	z	P>z	[95% C. I.]		X
Female	0.132	0.002	74.890	0.000	0.129	0.136	0.555
Age	0.002	0.001	1.920	0.055	0.000	0.005	16.757
Urban area	-0.086	0.003	-34.130	0.000	-0.091	-0.081	0.838
Mother's education	-0.031	0.001	-57.030	0.000	-0.032	-0.030	4.416
Monthly household income	0.046	0.001	38.190	0.000	0.043	0.048	5.970
Access to books in the home	-0.067	0.001	-58.440	0.000	-0.069	-0.064	1.746
Full-time schooling	-0.077	0.003	-27.740	0.000	-0.082	-0.072	0.133
Public status of the school	-0.267	0.007	-38.450	0.000	-0.281	-0.254	0.946
School tuition fees	-0.091	0.003	-32.510	0.000	-0.097	-0.086	1.178
log(Public spending per student)	-0.199	0.007	-26.780	0.000	-0.213	-0.184	14.943
% of teachers holding a postgraduate diploma	-0.005	0.000	-50.090	0.000	-0.005	-0.004	38.842

Panel B. Language

Variable	dy/dx	Std. Err.	z	P>z	[95% C. I.]		X
Female	0.003	0.002	1.630	0.103	-0.001	0.006	0.555
Age	0.002	0.001	1.370	0.170	-0.001	0.004	16.757
Urban area	-0.091	0.003	-36.190	0.000	-0.096	-0.086	0.838
Mother's education	-0.033	0.001	-61.720	0.000	-0.035	-0.032	4.416
Monthly household income	0.048	0.001	40.210	0.000	0.046	0.050	5.970
Access to books in the home	-0.065	0.001	-57.080	0.000	-0.067	-0.063	1.746
Full-time schooling	-0.058	0.003	-20.830	0.000	-0.063	-0.052	0.133
Public status of the school	-0.253	0.007	-35.180	0.000		-0.239	0.946
School tuition fees	-0.090	0.003	-31.630	0.000	-0.096	-0.085	1.178
log(Public spending per student)	-0.172	0.007	-23.320	0.000	-0.187	-0.158	14.943
% of teachers holding a postgraduate diploma	-0.003	0.000	-34.430	0.000	-0.003	-0.003	38.842

Note: dy/dx reflects a discrete change of dummy variable from 0 to 1.

73. As we showed in this section, there are several variables that can explain low performance in education, including policy variables. The findings are consistent with both higher public spending on education and better teacher qualifications playing a crucial role for improving academic performance and improving educational quality in Colombia.

Summary discussion and Conclusions

74. This paper investigated the links between higher public school spending, better teacher qualifications and student educational achievement. It finds that, for students from low-income families, higher per-student spending and better-qualified teachers is associated with higher performance in Mathematics and Language, after controlling for student and school characteristics such as family income,

gender, age, level of education of the mother, neighbourhood, and some school characteristics. These links are statistically significant and robust to a rich set of controls. For children from non-poor families, the links between educational achievement and higher school spending and improved teacher qualifications are weaker . The results highlight how improved access to school resources may profoundly shape the life outcomes of economically disadvantaged children, and may thereby significantly reduce the intergenerational transmission of poverty, provided that causality runs from policy variables to performance outcomes.

75. Suggestive benefit-cost analysis reveals that higher public investments in school spending are likely worthwhile. The results suggest that if all regions reach the level of spending per student of Bogota – the region with the highest spending – average math scores can increase by 3.8 to 4.3 points (around 8%), depending on the regions, with the highest improvement for low income students.

76. Given that public spending in education on Colombia has risen significantly since the 2000, the results of this paper might help explain the improved average performance of Colombian students in PISA exams, which was more important for children from low-income families than for students from high income families.

77. However, there is scope for continuing to improve education quality in Colombia. Increasing both public funding in education as well as improving teacher qualifications, particularly in the poorest departments, is likely to contribute in an efficient way to a better-skilled workforce, productivity gains and sustained economic growth.

REFERENCES

Avendaño, R., Barrera-Osorio, F., Nieto Parra, S. & Vever, F. (2016), Understanding student performance beyond traditional factors. *OECD Development Centre Working Papers*, No. 331. Paris: OECD Publishing. http://dx.doi.org/10.1787/5jlz7n7bxn0p-en

Bonilla, L. & Galvis, L. A. (2014), Profesionalización docente y calidad de la educación escolar en Colombia. En A. Sánchez, & A. Otero, Educación y desarrollo regional en Colombia (pp. 159-209). Bogotá: Banco de la República.

Bruns, B. & Luque, J. (2015), Great teachers: How to raise student learning in Latin America and the Caribbean. Washington, DC: World Bank.

Carlitz, R. (2009), Quantity Without Quality? Rethinking Education Goals. *Budget Brief*, Vol. 2(8). International Budget Partnership.

Chimombo, J. P. (2005), Quantity Versus Quality in Education: Case Studies in Malawi. *International Review of Education*, Vol. 55, pp. 155–172.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfield, F. D., & others. (1966), Equality of educational opportunity. Washington D.C.: U.S. Government Printing Office.

Duarte, J., Bos, M. S. & Martín Moreno, J. (2012), Calidad, igualdad y equidad en la educación colombiana (análisis de la prueba Saber 2009). Washington D.C.: Inter-American Development Bank.

García, M., Espinosa, J. R., Jiménez, F. & Parra, J. D. (2013), Capítulo 3. Los factores determinantes de la calidad de la educación: una revisión de literatura. En M. García, J. R. Espinosa, F. Jiménez, & J. D. Parra, *Separados y desiguales. Educación y clases sociales en Colombia* (pp. 32-81). Bogotá: Centro de Estudios de Derecho, Justicia y Sociedad, Dejusticia.

Gaviria, A. & Barrientos, J. H. (2001), Determinantes de la calidad de la educación en Colombia. Archivos de Economía. No. 159, noviembre. Departamento Nacional de Planeación.

Hanuskek, E. A. (2004), What if there are no 'best practices'? *Scottish Journal of Political Economy*, Vol. 51(2), pp. 156-172.

Hanushek, E. A. & Woessmann, L. (2015), The Knowledge Capital of Nations: Education and the Economics of Growth. Cambridge, MA: The MIT Press.

ICFES. (2015), Sistema Nacional de Evaluación Estandarizada de la Educación. Lineamientos generales para la presentación del examen de Estado SABER 11°. Bogotá: ICFES.

Lee, J-W. & Barro, R. J. (2001), Schooling quality in a cross-section of countries. *Economica*, Vol. 68, pp. 465-488.

OECD (2013a), PISA 2012 Results: What makes schools successful? Resources, policies and practices (Volume IV), PISA. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264201156-en

OECD (2013b), Teachers for the 21st Century: Using Evaluation to Improve Teaching. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264193864-en

OECD (2014), Education at a Glance 2014: OECD Indicators. Paris: OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en

OECD (2015), OECD Economic Surveys: Brazil 2015. Paris: OECD Publishing. http://dx.doi.org/10.1787/eco_surveys-bra-2015-en

OECD (2016a), Education in Colombia, Reviews of National Policies for Education. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264250604-en

OECD (2016b), Education at a Glance 2016: OECD Indicators. Paris: OECD Publishing. http://dx.doi.org/10.187/eag-2016-en

OECD (2016c), PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264266490-en

OECD (2016d), PISA 2015 Results (Volume II): Policies and Practices for Successful Schools, PISA. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264267510-en

Pritchett, L. (2006), "Does learning to add up add up? The returns to schooling in aggregate data", in E.A. Hanushek and F. Welch (eds.), Amsterdam: *Handbook of the Economics of Education*, pp. 635-695.

Sarmiento, A., Becerra, L., & González, J. I. (2000), La incidencia del plantel en el logro educativo del alumno y su relación con el nivel socioeconómico. *Coyuntura Social*, No. 22, Mayo, pp. 53-64.

STATISTICAL APPENDIX

Table A1. **Data definitions**

Variable	Definition	Source
Sex	Binary variable. Values: 1=female, 0=man.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
Age	Continuous variable. Values between 16 and 18 years old.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
School area	Binary variable. Values: 1=urban area, 0=rural area.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
Mother's education	Categorial variable. Values: 1=No education, 2=Incomplete primary education, 3=Complete primary educaton, 4=Incomplete secondary education, 5=Complete secondary education, 6=Incomplete technical professional and technological education, 7=Complete technical professional and technological education, 8=Incomplete tertiary education, 9=Complete tertiary education, 10=Postgraduate education.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
Monthly household income	Categorial variable. Values: 1=10 or more minimum wages, 2=Between 7 and lees than 10 minimum wages, 3=Between 5 and less than 7 minimum wages, 4=Between 3 and less than 5 minimum wages, 5=Between 2 and less than 3 minimum wages, 6=Between 1 and less than 2 minimum wages, 7=Less than the minimum wage.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
Access to books in the home	Categorial variable. Values: 1=Between 0 and 10 books, 2=Between 11 and 25 books, 3=Between 26 and 100 books, 4=More than 100 books.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
Full-time schooling	Binary variable. Values: 1=Full-time schooling, 0=No full-time schooling.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
Public status of the school	Binary variable. Values: 1=Public school, 0=Private school.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
School tuition fees	Categorial variable. Values: 1=No tuition fee, 2=Tuiton fees for less than 87,000 Colombian peso, 3=Tuition fees between 87,000 and less than 120,000 Colombian peso, 4=Tuition fees between 120,000 and less than 150,000 Colombian peso, 5=Tuition fees between 150,00 and less than 250,000 Colombian peso, 6=Tuition fees higher than 250,000 Colombian peso.	Saber 11 test, 2015 results (ICFES) (http://www.icfes.gov.co/investigadores-posgrado-2/acceso-a-bases-de-datos).
log(Public spending per student)	Continuous variable, at department or ETC level.	Ministry of Education (MEN).
% of teachers holding a postgraduate diploma	Continuous variable, at department or ETC level.	National Department of Statistics (DANE) (https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/poblacion-escolarizada/educacion-formal).

Table A2. **Descriptive statistics**

Variable	Observations	Mean	Standard Deviation	Min	Max
Mathematics score	373,353	50.64	11.453	0	100
Language score	373,353	50.269	8.937	0	100
Sex	371,746	0.554	0.497	0	1
Age	373,354	16.795	0.764	16	18
School area	342,337	0.838	0.369	0	1
Mother's education	372,976	4.495	2.03	1	10
Monthly household income	372,976	5.938	1.075	1	7
Access to books in the home	372,977	1.766	0.892	1	4
Full-time schooling	342,363	0.133	0.34	0	1
Public status of the school	373,354	0.868	0.339	0	1
School tuition fees	341,926	1.177	0.847	1	6
log(Public spending per student)	373,354	14.943	0.137	14.652	15.21
% of teachers holding postgraduate diploma	373,354	39.157	10.981	10.044	65.088

Table A3. **Correlation matrix**

Variable	Mathematics score	Language score	Sex	Age	School area	Mother's education	Monthly household income	Access to books in the home	Full-time schooling	Public status of the school	School tuition fees	log(Public spending per student)	% of teachers holding postgraduate diploma
Mathematics score	1.000	0.674	-0.157	0.010	0.118	0.301	-0.319	0.277	0.133	-0.175	0.248	0.115	0.043
Language score	0.674	1.000	-0.014	0.006	0.133	0.313	-0.323	0.279	0.116	-0.186	0.252	0.133	0.001
Sex	-0.157	-0.014	1.000	-0.006	0.014	-0.036	0.045	-0.013	-0.001	0.036	-0.024	-0.005	-0.003
Age	0.010	0.006	-0.006	1.000	-0.005	0.011	-0.036	0.013	0.019	-0.064	0.048	0.004	0.015
School area	0.118	0.133	0.014	-0.005	1.000	0.185	-0.160	0.127	-0.117	-0.035	0.008	0.178	-0.041
Mother's education	0.301	0.313	-0.036	0.011	0.185	1.000	-0.497	0.357	0.092	-0.311	0.361	0.137	-0.091
Monthly household income	-0.319	-0.323	0.045	-0.036	-0.160	-0.497	1.000	-0.422	-0.158	0.459	-0.555	-0.291	0.151
Access to books in the home	0.277	0.279	-0.013	0.013	0.127	0.357	-0.422	1.000	0.108	-0.204	0.258	0.219	-0.060
Full-time schooling	0.133	0.116	-0.001	0.019	-0.117	0.092	-0.158	0.108	1.000	-0.225	0.307	-0.031	0.012
Public status of the school	-0.175	-0.186	0.036	-0.064	-0.035	-0.311	0.459	-0.204	-0.225	1.000	-0.877	-0.199	0.153
School tuition fees	0.248	0.252	-0.024	0.048	0.008	0.361	-0.555	0.258	0.307	-0.877	1.000	0.208	-0.154
log(Public spending per student)	0.115	0.133	-0.005	0.004	0.178	0.137	-0.291	0.219	-0.031	-0.199	0.208	1.000	-0.382
% of teachers holding postgraduate diploma	0.043	0.001	-0.003	0.015	-0.041	-0.091	0.151	-0.060	0.012	0.153	-0.154	-0.382	1.000