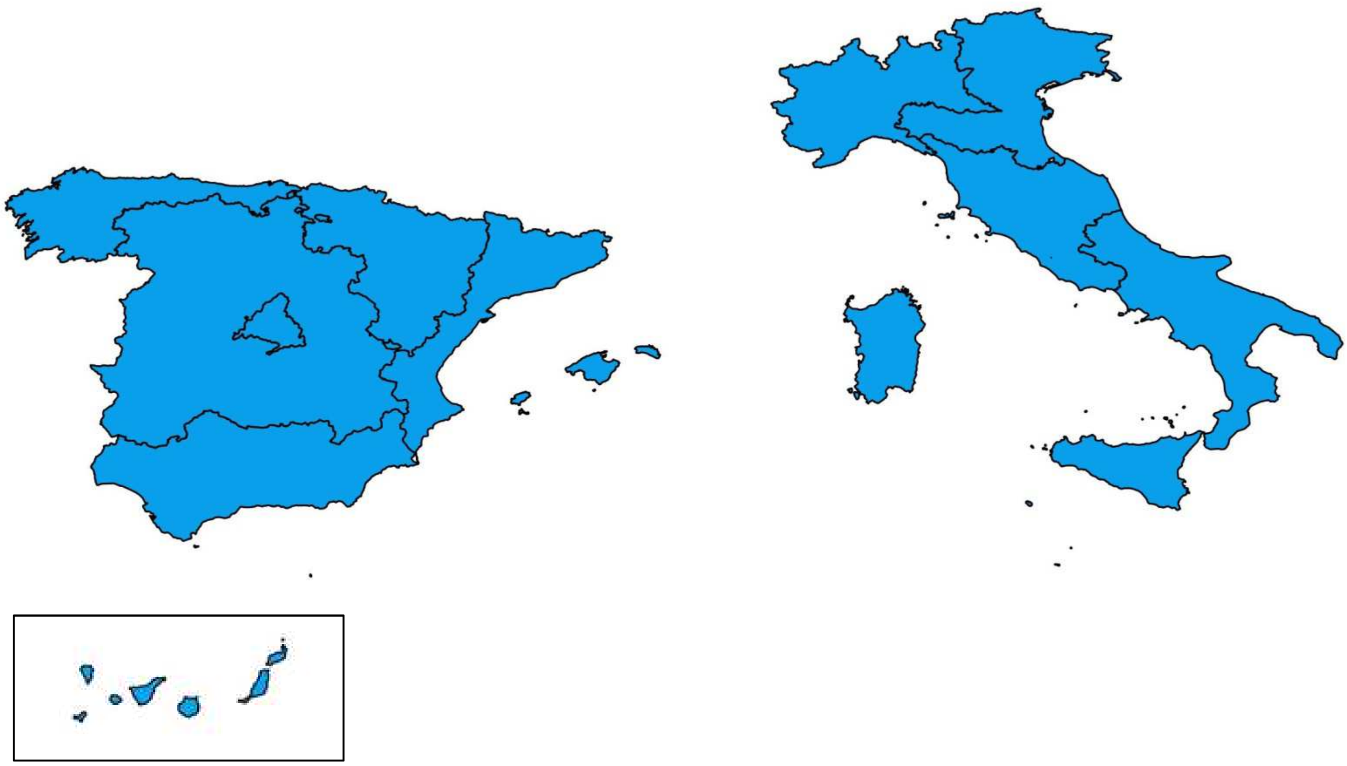
European Commission

JRC TECHNICAL REPORTS

# Regional inequalities in PISA: the case of Italy and Spain

Hippe, Ralph
Jakubowski, Maciej
Araújo, Luísa

2018



Joint Research Centre

EUR 28868 EN

**Contact information**
Name: Ralph Hippe
E-mail: ralph.hippe@ec.europa.eu

**JRC Science Hub**
https://ec.europa.eu/jrc

JRC109057

EUR 28868 EN

PDF ISBN 978-92-79-76296-3 ISSN 1831-9424 doi:10.2760/495702

Luxembourg: Publications Office of the European Union, 2018



How to cite this report: Hippe, R., Jakubowski, M. & Araújo, L. (2018). Regional inequalities in PISA: the case of Italy and Spain; Luxembourg (Luxembourg): Publications Office of the European Union; EUR 28868; doi:10.2760/495702.

# Contents

## Note

This report is part of the CRELL IX Administrative Agreement agreed between DG EDUCATION AND CULTURE (EAC) and DG JOINT RESEARCH CENTRE (JRC). Its content is stipulated under point 2.2 of the Technical Annex accompanying CRELL IX.

## Acknowledgements

The authors would like to thank colleagues from DG JRC and DG EAC for their useful comments on earlier versions of this report.

***Authors***

Hippe, Ralph (European Commission, JRC, Italy)
Jakubowski, Maciej (Warsaw University and Evidence Institute, Poland)
Araújo, Luísa (Instituto Superior de Educação e Ciências (ISEC), Portugal)

## Executive summary

This technical brief analyses the regional distribution of skills in Italy and Spain. Educational attainment rates have frequently been used as an indicator of regional educational development in EU Member States (MS). These rates indicate significant regional disparities in education within countries. However, recent evidence shows that the quality of education, as measured by the level of specific skills, is more important than the number of years one spends in school, in particular when considering the relationship between the cognitive (and non-cognitive) skills and economic growth. International large scale assessments (ILSA) of student performance measure these cognitive skills in key areas. OECD's Program for International Student Assessment (PISA) provides a very useful and important source of information of students' performance in key cognitive skills. When analysing PISA data, researchers and commentators often focus on cross-country comparisons. However, vast within-country differences exist, also in terms of educational attainment and PISA test scores. A focus on country averages alone would hence provide only a partial view of the status of education within countries. However, the possibility of exploring within-country differences with PISA data is limited to only a few countries. In this report we focus on regional inequalities in cognitive skills (as measured by PISA test scores) in Italy and Spain, using regional PISA data from the most recent 2015 wave, and we analyse the factors that are associated with these inequalities.

In order to insure full comparability between the two countries we define regions at the level of NUTS1 (macro-region), following Eurostat's official NUTS 2013 classification. We investigate regional inequalities by using descriptive statistics, by running a range of OLS regression models that allow us to analyse the associations between PISA 2015 science scores and the explanatory variables within regions and finally by using the Blinder-Oaxaca decomposition method to specify the factors that are related to within-country differences.

The results show that there are significant regional differences in PISA scores within both MS. There are several factors that are associated with regional differences within Italy and Spain. The factors most consistently positively associated with regional science achievement are teacher-directed teaching and epistemological beliefs, while grade repetition and truancy are significantly negatively related with achievement. Still, there is also a range of other relevant factors varying between and within both MS. The Blinder-Oaxaca decomposition also shows that variables such as the socio-economic background, the students' expected occupation, learning outside school time, truancy, immigrant status and grade repetition matter for within-country differences.

Our results suggest that policy makers should focus on finding solutions to limit truancy and rethink grade repetition to leverage scores in lower performing regions. Moreover, our results with regard to epistemological beliefs and teaching practices challenge thinking about how science should be taught in schools in Italy and Spain. The specific results for each region may allow policy makers to consider more in detail how a region stands in comparison to the rest of the country, and the specific factors that need to be addressed to improve the within-country inequality related to educational achievement.

# 1 Introduction

This technical brief pertains to DG EAC's strategic area of Efficiency and Effectiveness of Education Systems, and addresses the need to understand regional differences and draw policy implications related to regional inequalities.

The European Union has been facing rising regional economic inequalities in recent years (Financial Times, 2016). In particular, Italy and Spain are two Member States (MS) which are characterised by large regional disparities. For example, the OECD points out that "there is a 20 percentage point difference among unemployment rates between regions within Italy, Spain […], comparable to the difference between the national unemployment rate of Greece and that of Norway" (OECD, 2016a, p. 9). Differences among regions are also large in many other socio-economic areas (see OECD, 2016a). These facts make it necessary to better understand regional disparities also in the area of education, since it is the level of human capital that predominantly determines the economic development of a region (Gennaioli, La Porta, Lopez-de-Silanes, & Shleifer, 2013).[1] More specifically, a range of recent studies have shown that economic growth is most importantly affected by the quality of education and not by its quantity. That is, the cognitive skills of the population matter most (Hanushek & Wößmann, 2015). However, cognitive skills, as measured by PISA scores, may vary quite substantially within MS, in particular in Italy and Spain, as past PISA[2] cycles have already shown (see e.g., Matteucci & Mignani, 2014).

Therefore, this brief addresses the following questions: how large are regional inequalities in the most recent PISA 2015 round? Spain and Italy are two prominent examples providing PISA samples at the regional level, and this is why we focus on these two MS in this brief. In addition, what are the factors that are most strongly related to regional science scores and what are the variables that are associated to within-country disparities? In other words, we want to test whether various explanatory factors have a different relation with students' achievement in different regions. For example, have specific science teaching practices or grade repetition practices (which might change by region) varying effects in different regions? Since PISA 2015 includes a broader set of variables related to science teaching and learning than previous rounds, we are also able to consider a large range of possible relevant factors that may be related to regional inequality in science performance.

We investigate regional inequalities in Italy and Spain by using descriptive statistics, running a range of OLS regression models and decomposing within-country differences using the Blinder-Oaxaca method.

Our results show that there are large regional inequalities within Italy and Spain. When considering the determinants of students' performance, we find that the factors that – across all regions in both MS – are always significantly positively associated with PISA scores in our regional OLS models are epistemological beliefs and teacher-directed teaching, while truancy (skipping classes, etc.) and grade repetition are always negatively significant. Our Oaxaca-Blinder decomposition reveals that there is a range of different factors that account for within-country score gaps such as grade repetition, learning outside school time and socio-economic background (ESCS). Moreover, students' occupation expectations, truancy and immigrant status are other relevant factors.

These results may be considered by policy makers in shaping future education policies. For example, policies that improve truancy levels appear as effective devices for improving educational outcomes and thus to limit regional disparities in various cases. In addition, policy makers may reconsider how science is taught at school, as inquiry-based teaching is negatively and epistemological beliefs are positively related to PISA scores explaining within-country differences. Occupational expectations also play a role and

[1] See also e.g., Baten & Hippe (2017), Diebolt & Hippe (2018) and Hippe, Araújo, & Dinis da Costa (2016).

[2] PISA data are described more in detail in section 2. For more information see also the official OECD PISA website: http://www.oecd.org/pisa/.

their formation could be addressed by policies that inform students of their career options from an early age.

The technical brief has four main sections. First, we present the data and methodology. Second, we show the results of the empirical analysis. Third, we provide policy messages derived from our analyses. Finally, we suggest avenues for future research.

# 2 Data and methodology

## 2.1 Variables

Science achievement is the main domain in PISA 2015. This means that science performance was measured by a large number of science-related test items and every student taking the 2015 PISA test answered a number of science-related test items.[3] While reading and mathematics were also measured in 2015, the number of test items for these domains was smaller and the resulting measurement is less precise. Thus, we decided to focus on explaining regional differences in science performance.

Science is defined in PISA as "the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology, which requires the competencies to explain phenomena scientifically, evaluate and design scientific inquiry, and interpret data and evidence scientifically" (OECD, 2016b, p. 13). Student performance in science was measured using a computer-based test in both Italy and Spain. PISA provides also further information on students and schools. Data from student and school principal questionnaires can thus be used to analyse teaching and learning environments, school policies and the socio-economic background of the students' families.

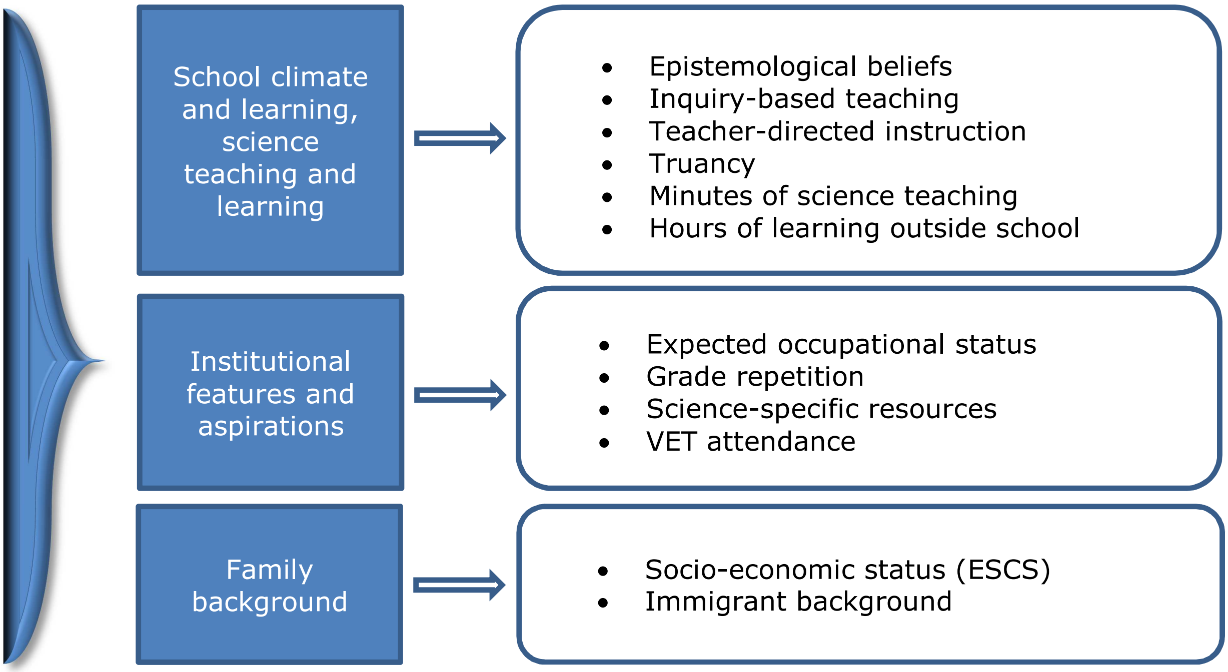
For the purpose of this brief we included most of the variables the OECD reports when presenting and discussing PISA 2015 cross-country differences and similarities (OECD, 2016b, 2016c, etc.). However, we exclude those variables that reflect national characteristics rather than regional ones.

Thus, the variables in our analyses can be categorised into 3 broad groups of predictors: (1) student responses about school climate and learning, science teaching and learning, (2) institutional features and aspirations (which could be affected by the geographic area of residence), and (3) family background variables (see Figure 1).[4]

[3] Student achievement is reflected by a set of plausible values that measure student outcomes and allow for estimating measurement error. In PISA 2015, ten plausible values are provided in the datasets and the analysis should replicate every estimation ten times with each plausible value. The results averaged across ten replications provide unbiased estimates of student achievement. In the calculation of standard errors variation across ten replications is included using special formulas that add estimates of measurement error to the estimates of sampling error. In this brief, all calculations for PISA 2015 data were performed using 10 plausible values and account for measurement error and sampling error by using clustered standard errors at the school level.

[4] See appendix for more details on the included variables.

**Figure 1.** Explanatory variables used in the analysis




The first group of variables refers to the learning environment, and specifically to approaches to teaching science, educational traditions and culturally-grounded educational practices and school climate. *Epistemological beliefs* reflect how students see science and scientific inquiry. Students who score higher in the PISA index of epistemological beliefs believe that "scientific knowledge is tentative (to the extent that students recognise that scientific theories are not absolute truths, but evolve over time) and adhere "to beliefs about the validity and limitations of empirical methods of inquiry as a source of knowing" (OECD, 2016c, pp. 99–100). Second, the instructional environment specific to the teaching and learning of science was operationalized in test items according to two constructs: inquiry-based teaching and teacher-directed instruction. The OECD-derived index of *inquiry-based science teaching and learning practices* refers to teaching methods where inquiry is at the core, while *teacher-directed instruction* focuses more on the role of the teacher (see appendix for further details).[5] PISA results indicate that an emphasis on inquiry-based teaching is negatively associated with student achievement. In contrast, the OECD-derived index capturing teacher-directed instruction has a positive relation with achievement (OECD, 2016d). In addition, *truancy* measures whether students are skipping school days, skipping classes, and arriving late at school. In fact, students were asked several questions about truancy. We used their responses to estimate an index of truancy using the principal component analysis based on polychoric correlations for ordinal responses (Kolenikov & Ángeles, 2004). The intuition is that skipping classes may have a negative association with performance. *Minutes of science teaching* is calculated by multiplication of teaching minutes in science with the number of classes per week. PISA 2015 results indicate that students score 5 points higher in science for every additional hour spent per week in science lessons (OECD, 2016d). *Hours of learning outside school* indicate the amount of time that students spend studying outside school. A high value might indicate a student who is having trouble with science learning or someone who aspires to have better results and looks for help outside school or works harder at home. At the regional level, the larger amount of learning outside school might reflect inefficiencies of the school system, which families compensate through private tutoring or teachers compensate through additional homework. However, it can also reflect a culture where more

[5] Note that these two teaching methods are not opposites, as a teacher can do experiments in the class and then spend time in class explaining students the context of the experiment and its outcomes in a teacher-directed way.

ambitious and richer families look for additional support outside school or students work harder studying at home. Thus, interpretation of associations between this variable and performance might vary depending on a country or a region.

*Expected occupational status*, or, in other words, students' aspirations, was measured through a question about expected profession that was coded into a specific expected occupational status index that was developed for international comparisons (see Ganzeboom, De Graaf, & Treiman, 1992). This index might reflect cultural, economic, social but also policy amenable factors. On the one hand, it might be expected that in more economically affluent regions students more often have ambitious future plans. However, higher student expectations might also reflect better teaching quality and policies that promote student ambitions and open ways to advance in education and later in career. *Grade repetition* indicates students who were repeating a grade before the age of 15. It controls for struggling students who are not promoted to next grade in their school system. Previous research with PISA data suggests that repeating a grade is not associated with improved student performance and shows a negative association with student attitudes (Ikeda & García, 2014). At the same time, this policy is very costly (Benhenda & Grenet, 2015). In PISA 2015, and in line with findings from previous PISA rounds, students who have repeated a grade at least once tend to have lower scores (OECD, 2016c). *Science-specific resources* are the resources a school has for science instruction. The index was constructed as a simple sum of school principal answers related to what resources she or he has available at school. *VET attendance* indicates whether students are in a vocational education track. The variable was obtained from the ISCEDO variable ("ISCED orientation") by coding as 0 students in general tracks or schools and as 1 all other students (those in pre-vocational or vocational tracks or schools). This follows the OECD's strategy to identify students in vocational education. It should be noted that in Spain most 15-year-olds are in academic or general schools, while vocational schools are more common in Italy. Thus, the results should be interpreted cautiously for Spain in this regard. Early tracking, before the age of 15, has been shown to increase educational inequality (Ruhose & Schwerdt, 2016).

*Economic, Social and Cultural Status (ESCS)* is an OECD index measuring student socio-economic background. PISA measures ESCS with an extensive set of questions related to parent occupation, education and household cultural, educational and economic resources. It is usually positively associated with PISA scores. *Immigrant background* refers to first and second generation immigrant students. In PISA 2015, a high concentration of immigrant students in a country is not associated with poorer student performance, while individually immigrant students on average show lower performance in European countries (OECD, 2016c).

## 2.2 Regions

In this study, regions are coded according to the NUTS (Nomenclature for Territorial Statistics) 2013 classification of the European Union. NUTS distinguishes various levels, ranging from NUTS0 (the country level) to NUTS3 (e.g., provinces in Spain).[6] The Eurostat classification of regions follows several specific criteria (see appendix for details). For standardization purposes, we merged available NUTS2 regional data into NUTS1 regions in the case of Spain, as the data for Italy are only available at this level (see Table 1). Thus, the regional unit of this study is NUTS1. In this way, we include a total of 12 regions, i.e., 5 for Italy and 7 for Spain.

[6] For more details, see appendix.

**Table 1.** Regional coverage

| MS | Overall NUTS coverage | National name for NUTS level | NUTS codes | Region name (national) | Region name (English) |
|---|---|---|---|---|---|
| ES | 7 NUTS1 | Agrupación de comunidades autónomas | ES1 | NOROESTE | NORTH-WEST |
| | | | ES2 | NORESTE | NORTH-EAST |
| | | | ES3 | COMUNIDAD DE MADRID | MADRID |
| | | | ES4 | CENTRO (ES) | CENTRE |
| | | | ES5 | ESTE | EAST |
| | | | ES6 | SUR | SOUTH |
| | | | ES7 | CANARIAS | CANARY ISLANDS |
| IT | 5 NUTS1 | Gruppi di regioni | ITC | NORD-OVEST | NORTH-WEST |
| | | | ITF | SUD | SOUTH |
| | | | ITG | ISOLE | ISLANDS |
| | | | ITH | NORD-EST | NORTH-EAST |
| | | | ITI | CENTRO (IT) | CENTRE |

## 2.3 Econometric models

In a first step, we run region-by-region OLS regressions on PISA science scores. These models allow us to see the variables that are significantly related to PISA scores within each region. We can also detect whether some variables have a higher relevance as concerns PISA scores in some regions than others.

However, using this approach we can only indirectly tackle the issue of within-country differences, as each region is dealt with separately. For this reason, in a second step we also decompose within-country differences using the Blinder-Oaxaca decomposition method. This method allows us to better understand the specific difference of a region with regard to the rest of the country, and to find out which category of variables and which individual variables are the most important factors for these disparities within countries.

# 3 Empirical analysis

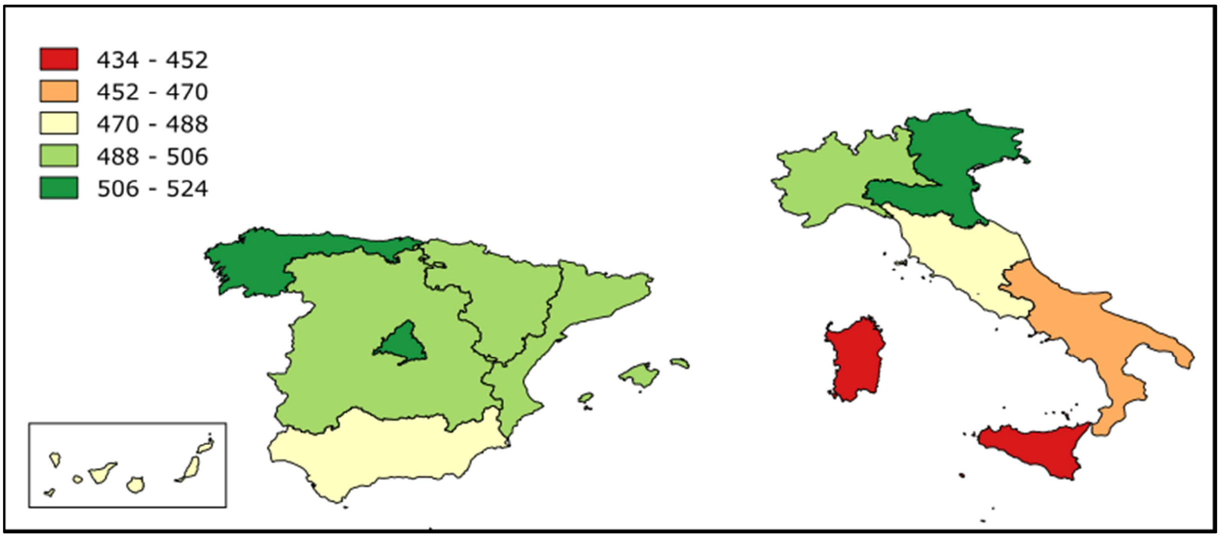
## 3.1 Descriptive statistics

In total, we have almost 44,000 student observations from both Italy and Spain in our dataset, of which roughly three quarters come from Spain.[7] Overall, the average science performance in Spain (494) is close to the OECD average (493), while it is significantly lower in Italy (480).

We begin by showing a PISA scores map to get a first intuition for the data (see Figure 2). As mentioned before, to increase regional homogeneity between both MS, we use NUTS1 regions for all our analyses. The map shows that there are high regional differences in Italy. They are also higher than in Spain, although Italy has a less decentralised education system.[8] In Italy the scores in the North-East are among the highest in Europe, whereas the scores in the Islands region are quite low.

[7] For more information, see descriptive statistics in the appendix.
[8] On the origins of Italy’s centralised school system, see Cappelli (2015, 2016).

**Figure 2.** Regional PISA 2015 science scores in Italy and Spain




Note: Classes of PISA scores defined according to equal intervals.

In fact, as also the descriptive statistics in Table 2 show, the gap among the Italian macro-regions is 90 score points, the equivalent of more than two years of schooling.[9] In Spain, there are also significant regional differences but they are less pronounced, with differences equivalent to more than one year of schooling. In Spain, the top performer at the NUTS1 level is the capital region, Madrid, followed by the North-west region, while the lowest scores are recorded for the South and the Canary Islands.

**Table 2.** Descriptive statistics for regional PISA 2015 science scores in Italy and Spain

| NUTS code | Region name | Mean | S.E. | 95% confidence interval | |
|---|---|---|---|---|---|
| | | | | Lower limit | Upper limit |
| ES1 | NORTH-WEST | 506.09 | 1.92 | 502.26 | 509.92 |
| ES2 | NORTH-EAST | 494.01 | 1.82 | 490.40 | 497.63 |
| ES3 | MADRID | 516.17 | 3.30 | 509.62 | 522.73 |
| ES4 | CENTRE | 500.69 | 1.82 | 497.06 | 504.32 |
| ES5 | EAST | 499.48 | 2.79 | 493.93 | 505.02 |
| ES6 | SOUTH | 472.63 | 3.51 | 465.65 | 479.62 |
| ES7 | CANARY ISLANDS | 473.82 | 3.00 | 467.84 | 479.80 |
| ITC | NORTH-WEST | 497.92 | 4.46 | 489.05 | 506.79 |
| ITF | SOUTH | 458.79 | 3.60 | 451.62 | 465.97 |
| ITG | ISLANDS | 433.82 | 8.03 | 417.83 | 449.80 |
| ITH | NORTH-EAST | 523.93 | 5.16 | 513.66 | 534.21 |
| ITI | CENTRE | 481.41 | 7.09 | 467.29 | 495.53 |

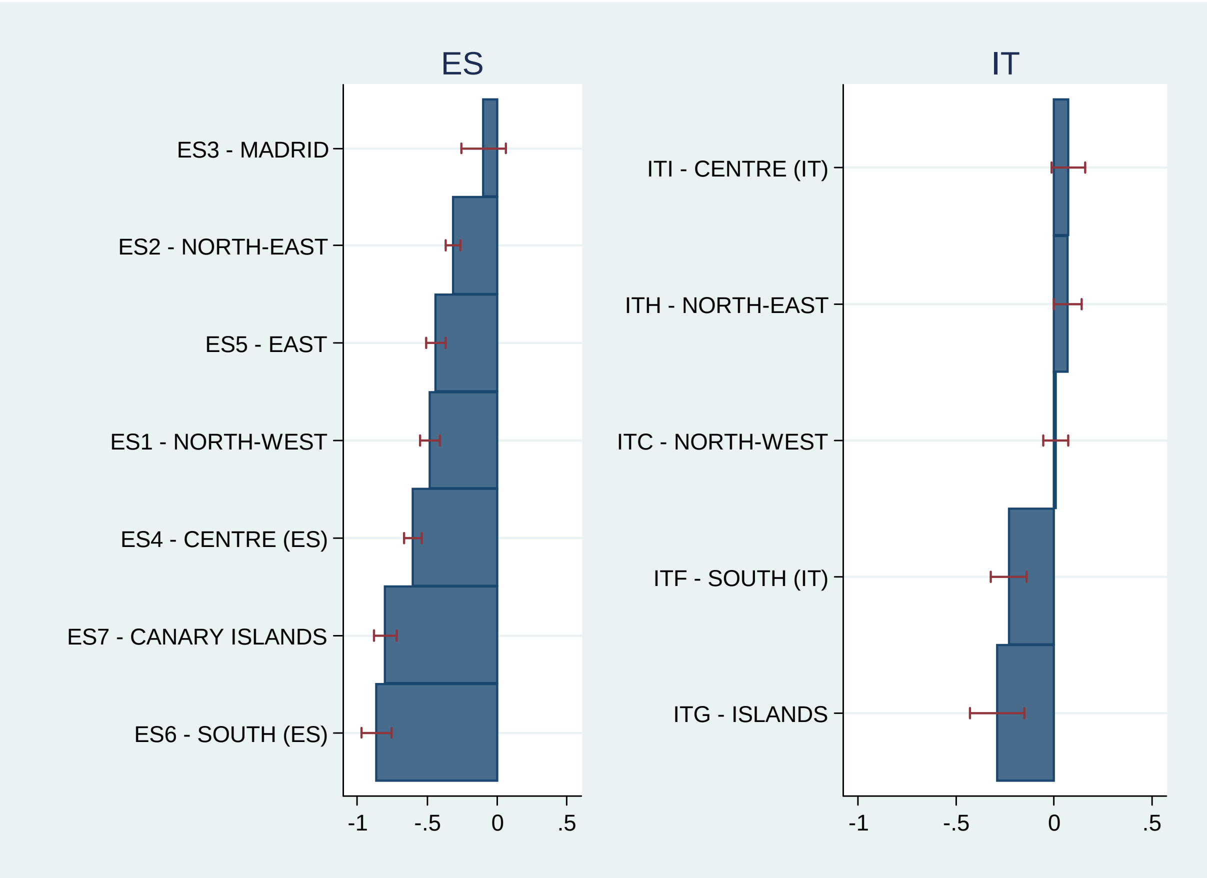
Next, we present in detail a number of explanatory variables in our analysis: the ESCS, immigrant background, epistemological beliefs, truancy, inquiry-based science teaching, students' occupational expectations, learning time outside school and grade repetition.[10] Our later econometric analyses will show that these are particularly relevant (i.e., significant) factors. Apart from mean values per region, we also show the 95 %

[9] See similar interpretations of score point differences in OECD (2010, p. 30) and OECD (2016e, p. 7).
[10] See the appendix for the illustration of additional variables in our analysis and the descriptive statistics for all variables.

confidence intervals. For the categorical variables immigrant background and grade repetition, we also compute score points for each group to provide further comparative information.

First, we show the distribution of the average values of the ESCS at the regional level for both Italy and Spain. We see a high variance in the average ESCS among Italian (IT) regions (see Figure 3). More specifically, the two southern/Island regions have a substantial negative average ESCS (relative to an OECD average of 0), while they have positive values for the North-east and the Centre.[11] In addition, some of the lowest average values of the ESCS – also in a European regional comparison – are present in Spain (ES; in particular in the South and Canary Islands). In fact, even the region with the highest ESCS (Madrid) is below the OECD average of 0. Regional differences are even higher in this MS than in Italy.

**Figure 3.** Regional averages for ESCS in Italy and Spain



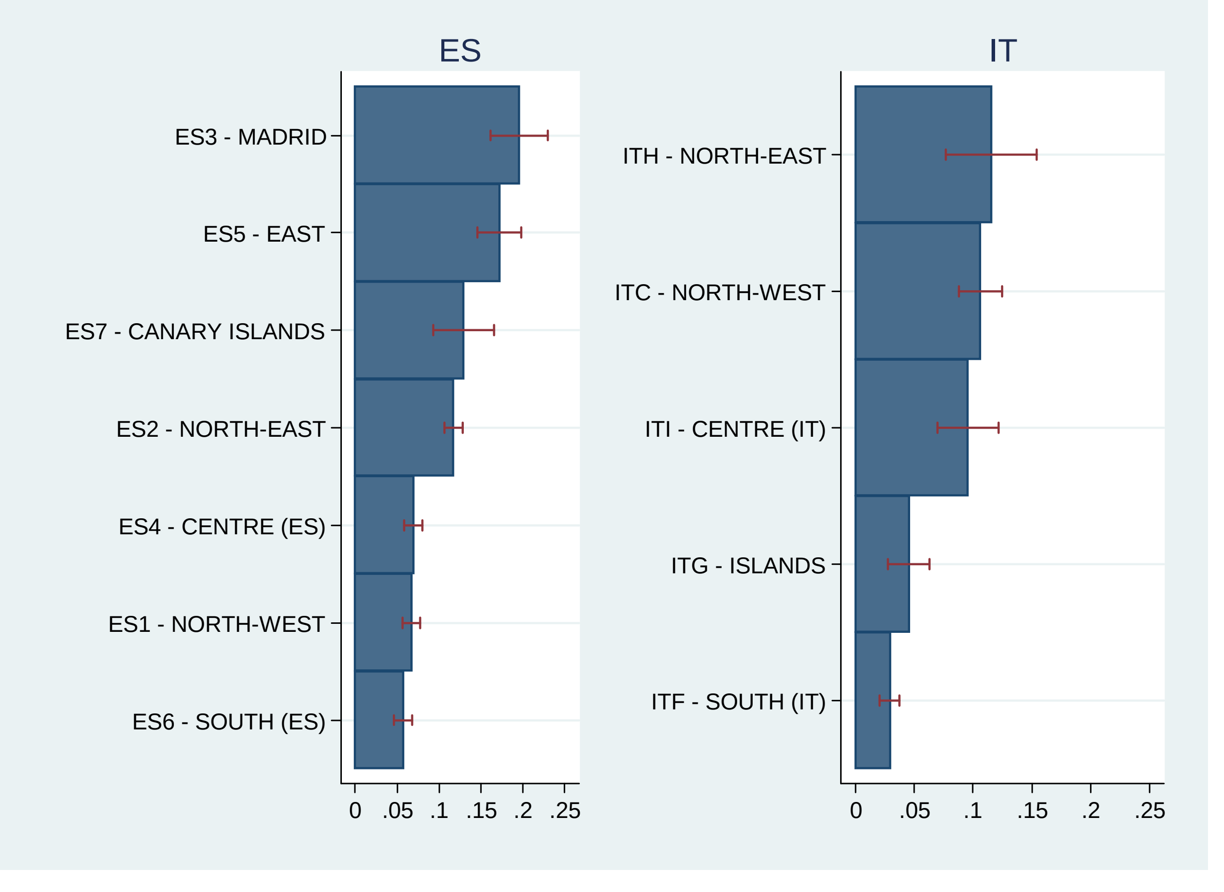

[11] The ESCS is a z-score variable as all OECD-derived indices. This means that it has a mean of 0 and a standard deviation of 1 across OECD countries with weighting each country equally. The same is true for other indices that we are using as, for example, epistemological beliefs, inquiry-based teaching and truancy.

Second, immigrant students are present in all Italian and Spanish regions (see Figure 4). However, their share is quite unevenly distributed. In both MS, immigrants may select their destinations primarily according to the attractiveness of the local working or living conditions, so that in Spain the highest shares of immigrant students are found in Madrid and the East, while in Italy they are in both Northern regions. In both MS, the lowest share of immigrant students is recorded in the Southern regions.

In addition, immigrant students always score lower than natives in all regions in Italy and Spain (see appendix). In Spain, the score difference between natives and migrants is smallest in the Canary Islands, while it is highest in the North-east and in Madrid. In the North-east, this finding can be explained by the rather low scores of migrants, which are on a similar low level in the South (where migrants score lowest). In Italy, there are only minor differences between these two groups in the South (the lowest of all regions in Italy and Spain), while the gap is more than 70 score points in the North-east. The large difference in the North-east is not due to a low performance of migrants, as the migrants in this region perform similar to migrants in most other regions (except in the Islands region), but is the result of the exceptional high performance of native students in this region. Thus, migrants show a homogeneous performance across most Italian regions (always between 450 and 460 score points) – except in the Islands region, where their scores are at a lower level, but natives also score lower here than elsewhere. In contrast, the scores of natives vary much more across regions than do those of migrants.

In general, the rough tendency in both MS is that the higher overall score of the region, the higher are also the differences between native and migrant students. This observation, however, does not always hold for all regions.

**Figure 4.** Regional averages for immigrant background in Italy and Spain



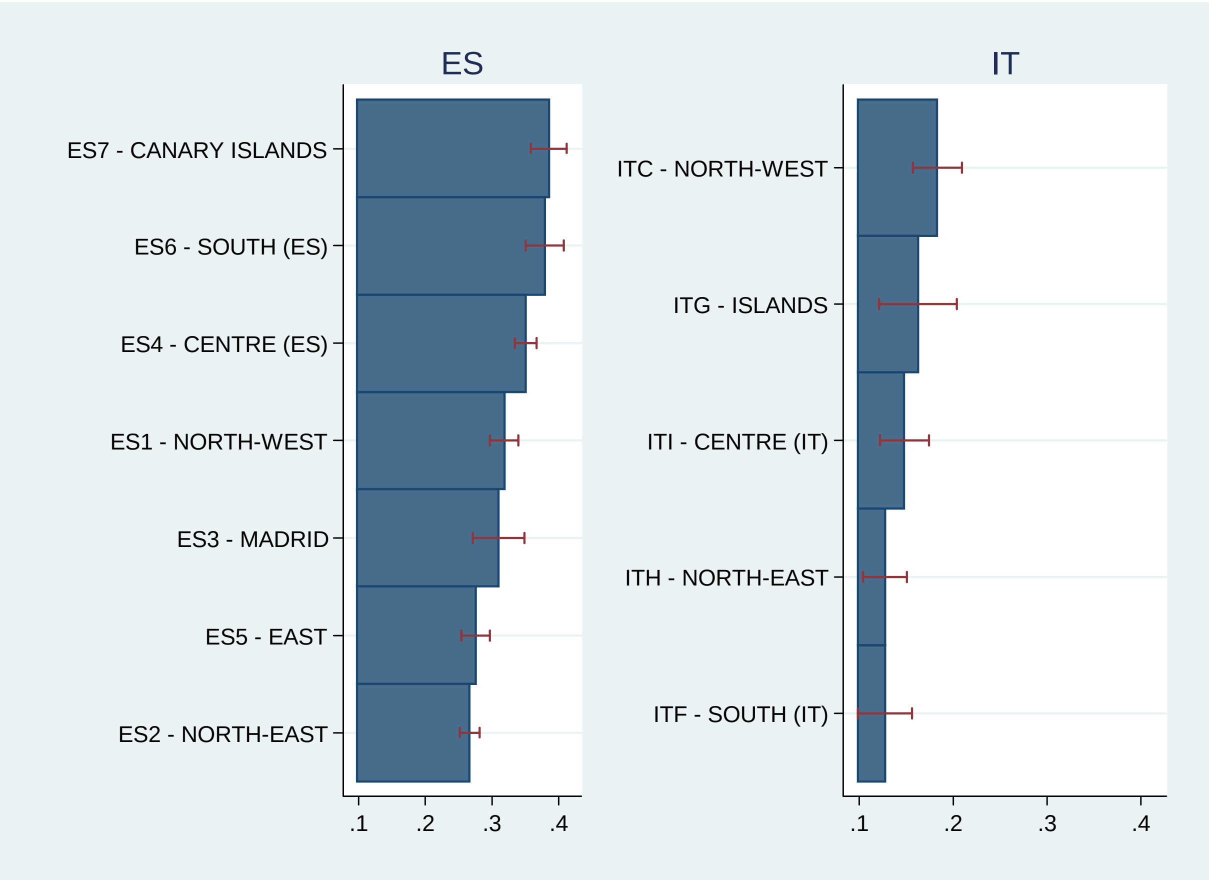
Grade repetition practices are clearly different in both MS, as the levels in Spain are always much higher than in Italy. Still, there is a substantial variation also within these countries, as shows Figure 5. The geographical pattern is also quite different: while in Spain the highest rates of grade repetition are recorded for the Southern part of the country (Canary Islands, South), in Italy the South region has the lowest grade repetition, while the Islands regions has one of the highest, only surpassed by the North-west.

Moreover, while grade repetition is more common in some regions in Spain than in others, the score point differences between those students who have repeated a grade and those who have not are always very large (see appendix): they mostly account for two and a half years of schooling (around 100 score points). Only in the East the difference is a bit smaller, but still very high (86 points). Grade repeaters, like non-repeaters, score lowest in the South and Canary Islands, while they score highest in the Madrid and the North-west regions – in other words, their performance is in line with the overall performance of each region.

The same result applies also to Italy. Still, while in Italy grade repetition is less common than in Spain, the performance difference of those who do repeat a grade is significantly smaller, but still large: it ranges between 60 to 86 points. Interestingly, the highest score difference is found in the Centre and South regions, while the smallest gap exists in the North-west.

Thus, in both Italy and Spain the score point differences between the two groups do not seem to be directly related to the overall performance of the region.

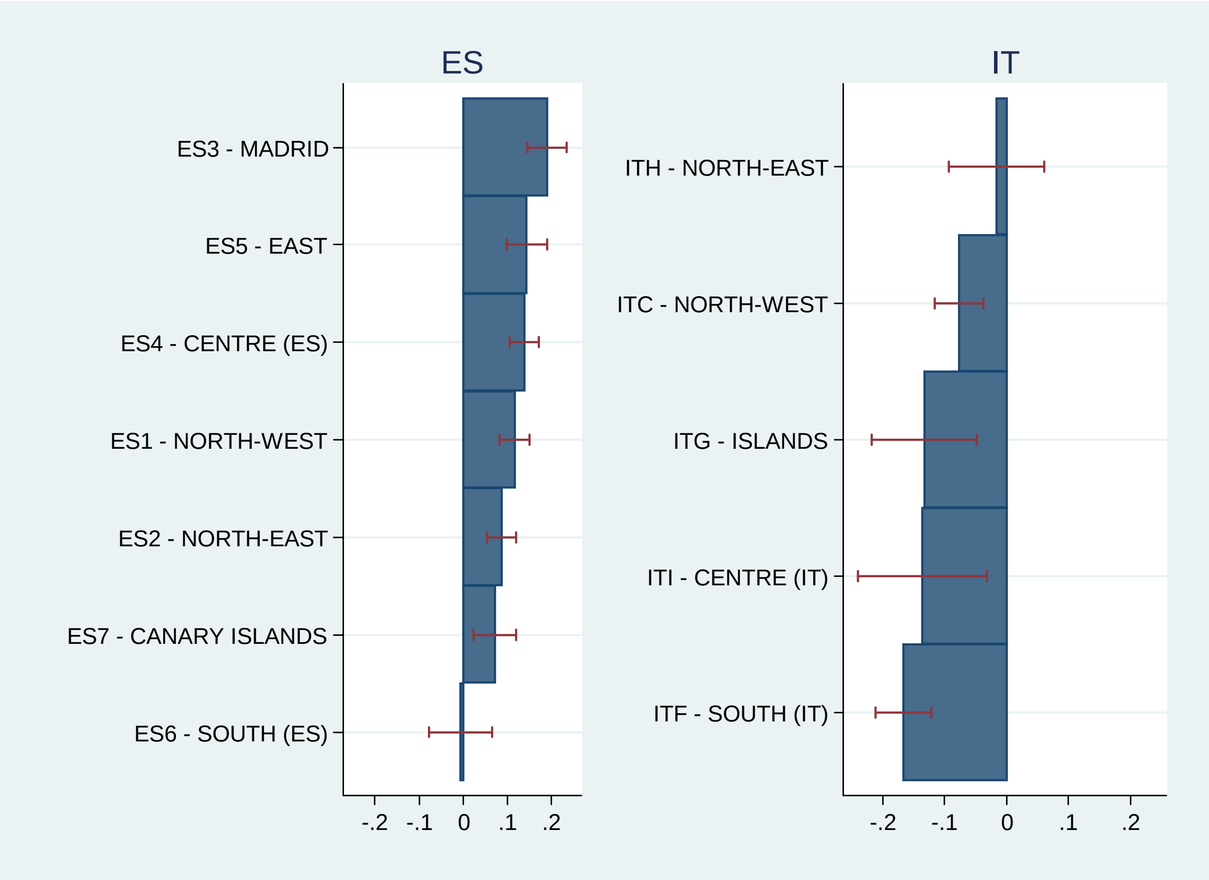
**Figure 5.** Regional averages for grade repetition in Italy and Spain



In addition, Figure 6 shows that North-east Italy has the highest values in epistemological beliefs, while the South has the lowest ones. Interestingly, the average values for all regions are negative, which indicates that they are lower than the OECD average (i.e., 0). This may indicate that there are substantial possibilities of improvement in this area, for instance by communicating to teachers the importance of teaching students that science concepts evolve over time and are subject to change and empirical confirmation.

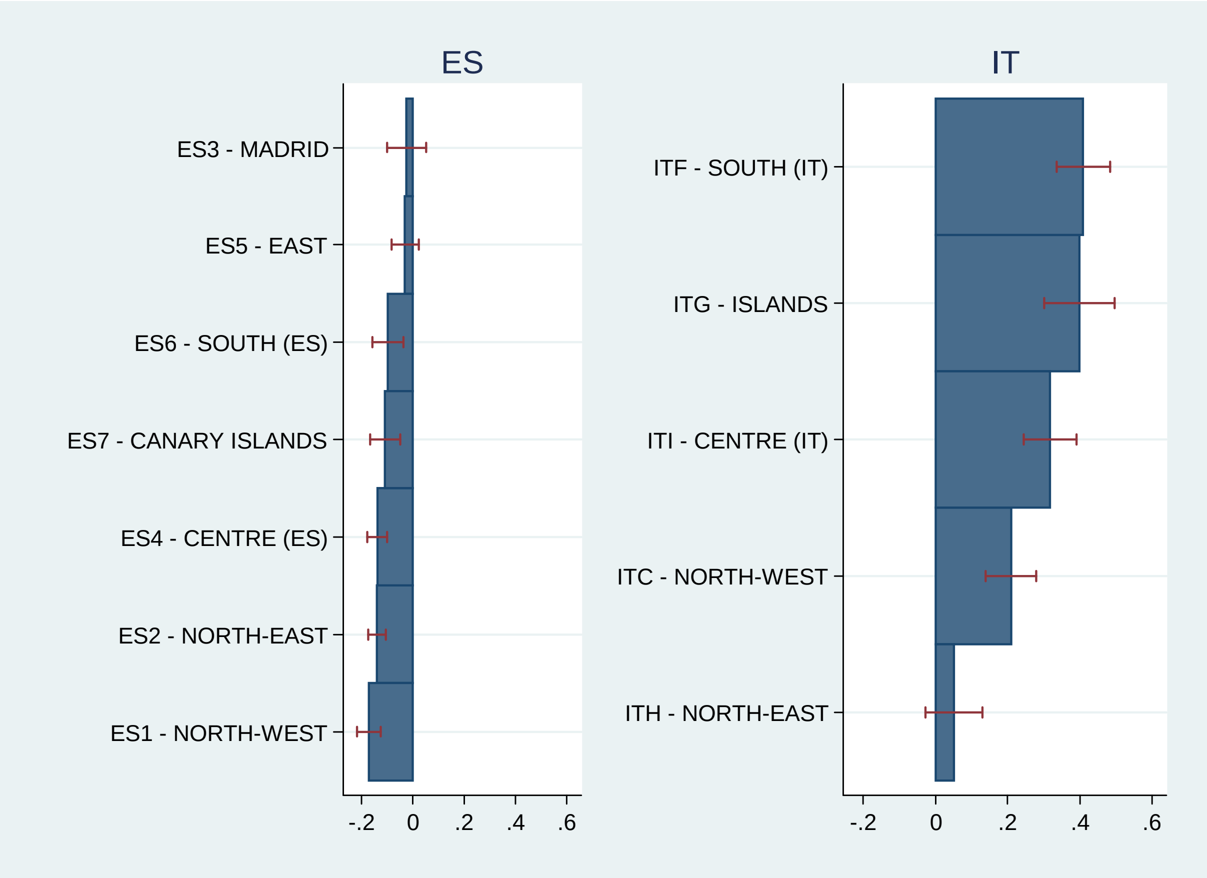
In Spain, only the South has a slight negative value, whereas all other regions have epistemological beliefs above OECD-average. As in the case of the other preceding variables, Madrid takes the lead as the Spanish region with higher epistemological beliefs.

**Figure 6.** Regional averages for epistemological beliefs in Italy and Spain



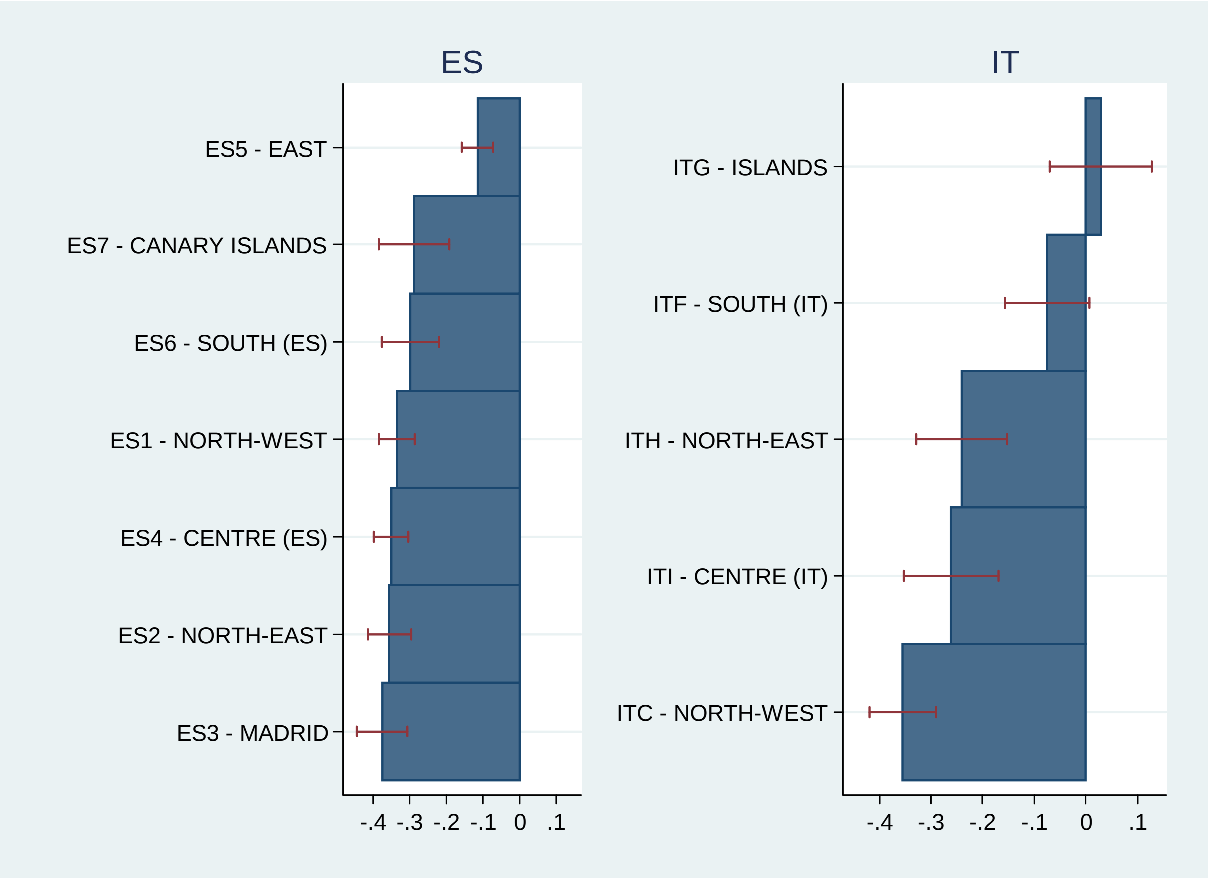
Similarly, we find again an opposing picture between Italy and Spain when considering truancy (see Figure 7). Truancy levels in some Italian regions (South, Islands, Centre) are among the highest in Europe, while the North-east has much lower but still above-OECD average levels. In contrast, in Spain all regions have below-OECD average truancy levels (i.e., below 0), which are indicated by their negative values. Spain's North-west features the lowest levels, while the values in Madrid are close to the OECD average.

**Figure 7.** Regional averages for truancy in Italy and Spain



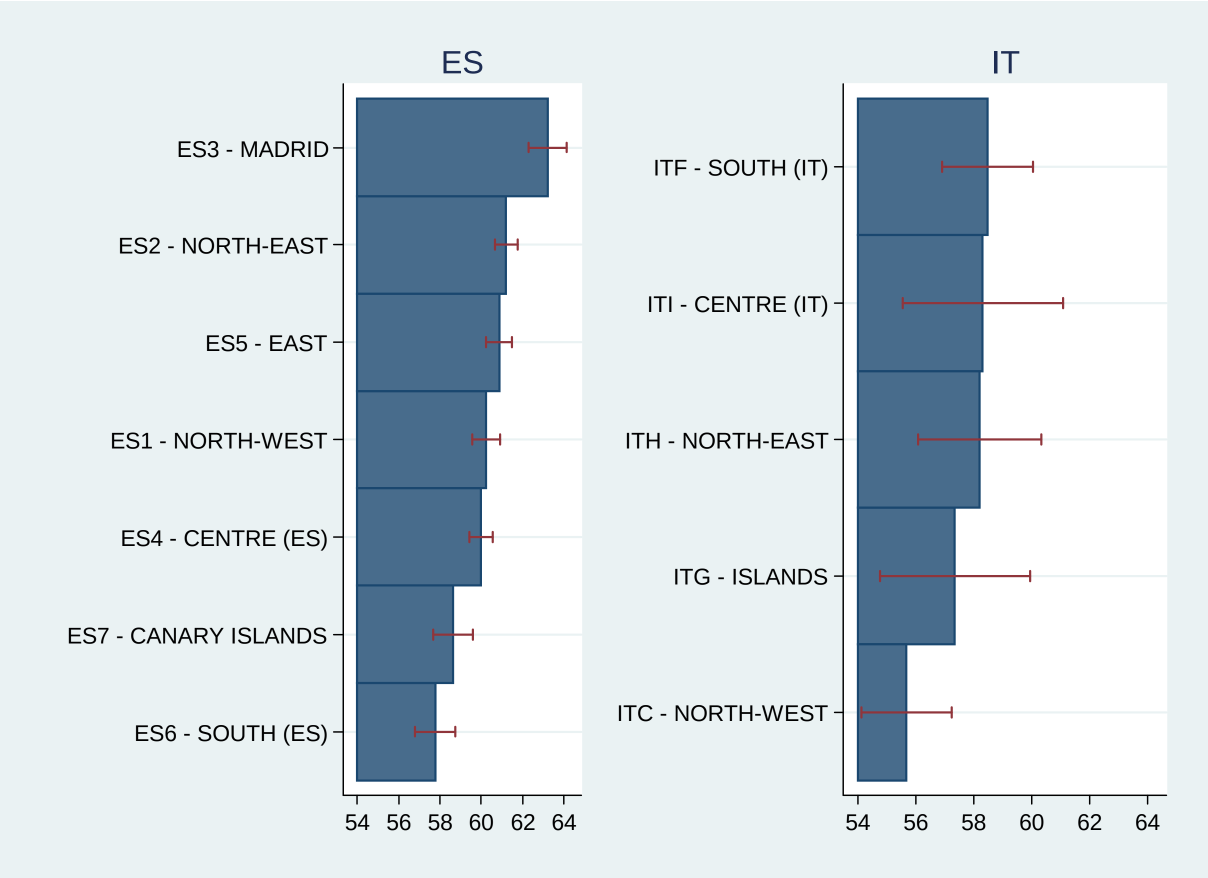
Moreover, inquiry-based teaching varies quite substantially among the regions (see Figure 8). The highest negative value for this variable is found in Spain for Madrid, while the East distinguishes itself from the other regions by its relatively high (but still below OECD average and hence negative) level. In Italy, the differences are even more pronounced, with the Islands region having (as the only region in the sample) a positive value, while the North-west is at the lower end of the spectrum.

**Figure 8.** Regional averages for inquiry-based teaching in Italy and Spain



Another relevant variable is the students' expected occupational status (SEI; see Figure 9). In Spain, we see that the lowest performing South and Islands regions also have the lowest expectations, while they are the highest among the top-performers Madrid and North-west. This rough relationship with scores is not evident in Italy. Here, the low-performing South region has the highest values for expectations, while the high-performing North-west is at the lower end.

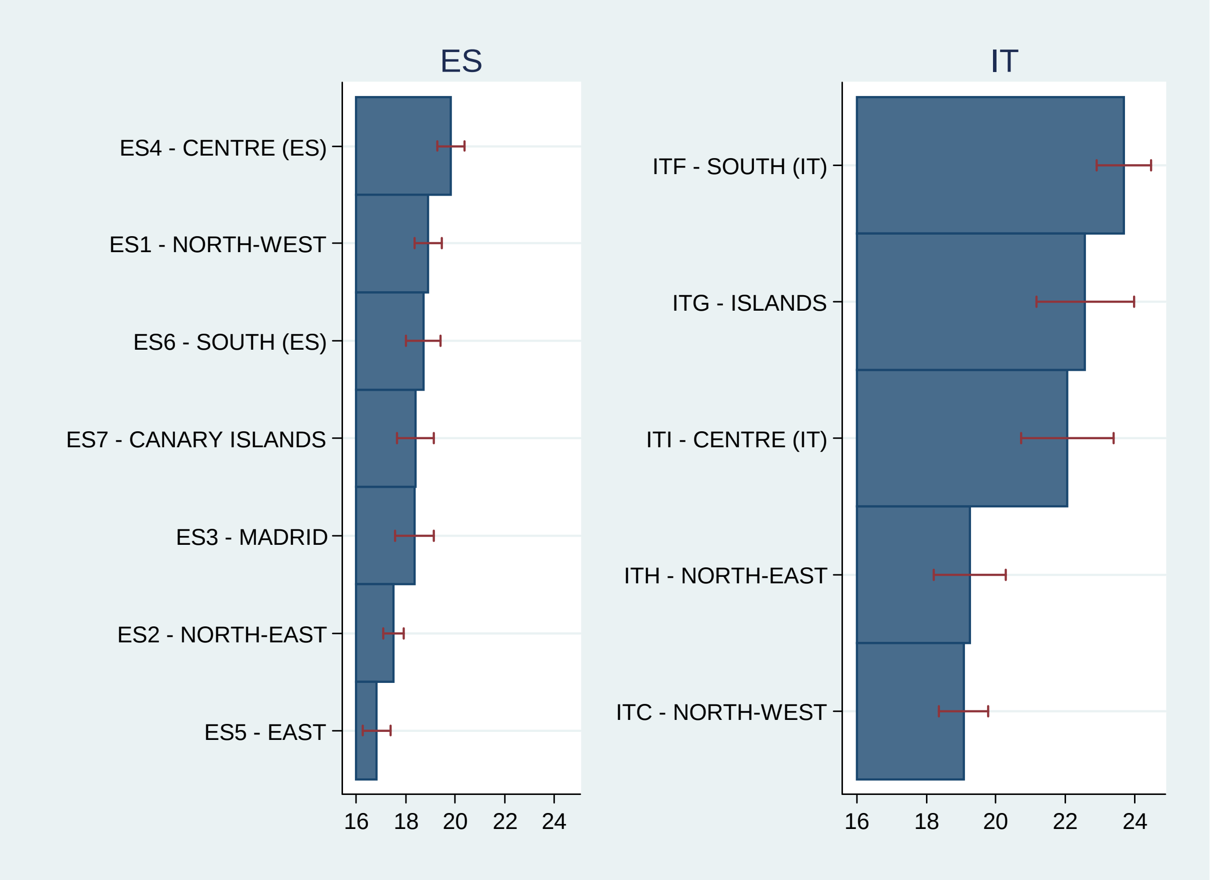
**Figure 9.** Regional averages for students' expected occupational status in Italy and Spain



Learning outside school is done on average for longer hours in Italy than in Spain, as Figure 10 shows. In general, learning outside school time can be related to the fact that students need to repeat once again the science content taught at school in their free time using a private tutor, or to the fact that they spend time doing their homework. Thus, this variable may indicate either an inefficient school system or particularly ambitious students.

In Italy, the North-South divide becomes once again apparent: the students with the lowest average number of outside school hours live in the North regions, while those with the highest levels are located in the South/Island regions. In Spain, the geographical distribution is not so clear, with the East having the lowest and the Central region having the highest average number of hours.

**Figure 10.** Regional averages for learning outside school in Italy and Spain




In a nutshell, we see that there is high variance for all the presented variables among the regions in Italy and Spain. This shows that the regional level offers interesting insights for understanding better educational achievement levels, which we address next by using regression analysis. In addition, while both Italy and Spain share many socio-economic and cultural bonds, we found in a number of cases different values among the regions of both MS, suggesting that the education systems face quite different challenges regarding the regional student population.

## 3.2 Regression analysis

### 3.2.1 OLS models

We estimated OLS models region-by-region, regressing the PISA scores on the previously indicated set of student-level variables. This allows analysing the association of our explanatory variables with PISA scores at the regional level.

We consider each MS separately. The data for Spain (see Table 2) show that all variables, except science-specific resources, VET and immigrant background are significant in all regions. On the one hand, a higher level of truancy, more learning outside school, more inquiry based (IB) teaching and having repeated a grade, attending a VET track and having an immigrant background decrease students' scores. On the other hand, higher epistemological beliefs, more teacher directed (TD) teaching, higher occupational expectations (SEI) and a higher socio-economic status (ESCS) increase scores in all regions. In all regions, science-specific resources are not associated with performance.

Furthermore, grade repetition is strongly associated with science scores in all regions, indicating a much lower performance of students repeating grades. Considering the association of truancy with science performance, for example, the most negative association is found in Madrid (ES3, -9.03), while it is significantly smaller in the Canary Islands (ES7, -4.47). Similarly, the ESCS plays a significantly stronger role in the South (ES6, 10.34) than in the North-west (ES1, 3.80).

**Table 3.** Regional OLS regressions for Spain

| | ES1 NORTH-WEST | ES2 NORTH-EAST | ES3 MADRID | ES4 CENTRE | ES5 EAST | ES6 SOUTH | ES7 CANARY ISLANDS |
|---|---|---|---|---|---|---|---|
| Truancy | -7.69*** | -6.30*** | -9.03*** | -5.24*** | -7.30*** | -6.90*** | -4.47* |
| Epist. beliefs | 16.06*** | 19.00*** | 15.85*** | 16.78*** | 14.60*** | 15.59*** | 17.82*** |
| Learning outside school | -0.52*** | -0.33*** | -0.41** | -0.21* | -0.41*** | -0.25* | -0.45** |
| IB teaching | -5.55*** | -6.64*** | -5.82* | -7.46*** | -5.21*** | -6.62*** | -8.79*** |
| TD teaching | 4.38** | 5.66*** | 8.20** | 6.84*** | 5.55** | 8.69*** | 6.60** |
| Min. of science teaching | 0.10*** | 0.10*** | 0.10*** | 0.12*** | 0.14*** | 0.09*** | 0.11*** |
| Science-spec. resources | -1.07 | 1.15 | -0.19 | 0.13 | 0.16 | 1.71 | 0.89 |
| SEI | 0.92*** | 0.66*** | 0.74*** | 0.63*** | 0.75*** | 0.60*** | 0.42*** |
| Grade repetition | -73.71*** | -61.52*** | -72.08*** | -70.32*** | -57.42*** | -75.11*** | -76.49*** |
| VET | -52.82*** | -15.95 | -55.80* | -30.46*** | -5.37 | -17.69 | -17.89 |
| ESCS | 3.80** | 6.76*** | 8.04*** | 6.88*** | 10.84*** | 10.34*** | 8.73*** |
| Immigrant background | -17.48*** | -19.68*** | -18.89*** | -8.56 | -24.74*** | -16.35* | -7.54 |
| Constant | 470.53*** | 453.01*** | 477.19*** | 468.82*** | 454.83*** | 455.49*** | 469.20*** |
| Number of students | 5579 | 8745 | 1808 | 5556 | 5191 | 3609 | 1842 |

*Note:* Significance level: * p<0.05, ** p<0.01, *** p<0.001.

Next, we take a closer look at the Italian regions (see Table 3). Some variables that we have found to be significant in the Spanish case are also very relevant in Italy. Overall, however, the picture appears to be more diverse in this MS. The variables that are always positively significant in all Italian regions are epistemological beliefs and TD teaching, while truancy and grade repetition are always negatively significant. Interestingly, all our variables are significant in the North-west (ITC), while this is not true for the other regions, indicating that our model fits to varying degrees the regions.

On a general note, we want to point out that some variables are more strongly related to PISA scores in one region than in others. For example, a unit increase in epistemological beliefs in North-west Italy (ITC) translates into a decrease of PISA scores by 16.32, while the coefficient is significantly higher in the case of North-east Italy (ITH, 24.19) suggesting that this factor plays a more important role in explaining student achievement in the latter. These results show that the 'classical' North-South divide in Italy does not always hold in our analyses. In addition, immigrant background does only have a negative significant association in northern Italy, which is the area with the highest immigrant student population.

**Table 4.** Regional OLS regressions for Italy

| | ITC NORTH-WEST | ITF SOUTH | ITG ISLANDS | ITH NORTH-EAST | ITI CENTRE |
|---|---|---|---|---|---|
| Truancy | -9.94*** | -9.94*** | -9.91*** | -7.26*** | -8.00* |
| Epist. beliefs | 16.32*** | 19.68*** | 20.13*** | 24.19*** | 24.00*** |
| Learning outside school | -0.48*** | -0.02 | -0.15 | -0.56** | -0.18 |
| IB teaching | -10.76*** | -10.38*** | -16.32*** | -0.79 | -8.72** |
| TD teaching | 15.57*** | 13.57*** | 14.72*** | 11.70*** | 14.47*** |
| Min. of science teaching | 0.08** | 0.09*** | 0 | 0.04 | 0.07 |
| Science-spec. resources | 7.38*** | 1.72 | 8.44* | 6.16** | 7.16*** |
| SEI | 0.70*** | 0.73*** | 0.88*** | 0.50** | 0.41 |
| Grade repetition | -29.47*** | -36.08*** | -34.38** | -35.16*** | -37.74*** |
| VET | -27.84** | -30.77*** | -20.37 | -1.41 | -40.27** |
| ESCS | 9.89*** | 9.33*** | 4.49 | 11.25*** | 11.34*** |
| Immigrant background | -23.76*** | 3.87 | -24.01 | -42.26*** | -11.05 |
| Constant | 447.52*** | 426.61*** | 375.48*** | 479.76*** | 447.40*** |
| Number of students | 2410 | 2324 | 807 | 5111 | 931 |

*Note:* Significance level: * p<0.05, ** p<0.01, *** p<0.001.

In sum, the results indicate that in each MS there are different variables that have an important role in relation to PISA science scores. Their importance also varies according to the considered region, sometimes substantially.

### 3.2.2 Blinder-Oaxaca decomposition

In this section we aim at identifying the variables that are significantly related to the within-country *differences* in PISA science scores, both in terms of the levels of these variables and in terms of the association between each variable and student performance.

To this end, we estimate Blinder-Oaxaca decomposition models (Blinder, 1973; Oaxaca, 1973) for all regions in Italy and Spain separately, comparing in each case one region's performance to that of the remaining regions in the same country. This approach enables us to understand how the differences in PISA science test scores between each region and the rest of the country are related to the variables in our model. More specifically, we run a set of regressions estimating the association between our set of explanatory variables and PISA scores for two groups:

- Group 1: Only the considered region
- Group 2: Rest of the country (i.e., all other regions of the country)

Based on these regression results, the Oaxaca-Blinder method decomposes performance differences across regions in each country into the endowment (corresponding to the level of a variable) and the coefficient effects (which can be interpreted as the "return" to the endowment). The basic intuition of this method is the following. The observed differences in the outcome variable (PISA test scores) between one region (Group 1) and all the remaining regions considered together (Group 2) are decomposed into two main components. The first one, corresponding to the endowment effects, captures the effect of the different distributions of the drivers (the endowments) between the two groups. It hence reflects the variability in the explanatory variables between the two groups. The second effect captures the effect of the differences in the coefficients (i.e. the returns) on the explanatory variables.

A simple example might help. Let us imagine that PISA scores just depend on the socio-economic status (ESCS) and the immigrant background. Then, the observed differences in PISA scores between Group 1 and Group 2 would be decomposed into i) an "endowment" effect, reflecting how the two groups differ in terms of ESCS and prevalence of immigrants; ii) a "return" effect capturing the variation in the coefficients on ESCS and immigrant status between the two regions.

We also report results for the "interaction" terms that estimate the combined effect of endowments and coefficients in some regions. Table 4 provides detailed results of this decomposition, reporting also the results separately for each variable.[12] These results help to understand which specific factors characterize the scores of each region in comparison with the rest of the country. This, in turn, might be useful when drawing policy implications.

The upper part of Table 5 (marked as "I" on the right-hand side) reports first the mean score predictions for each group and then their difference. In the case of ES1 (Galicia, Asturias and Cantabria, i.e. North-west Spain), the mean of PISA science scores is 507 for this region and 493 for the rest of Spain, yielding a significant gap of 14 PISA score points between this region's score and the rest of the country. In other words, the students of the region ES1 performed significantly better than those in the rest of the country. The difference is even larger for ES3 (Madrid), while it is smaller but still significant for ES4 (Centre) and ES5 (East). In contrast, ES6 (South) and ES7 (Canary Islands) have significantly lower scores compared to the rest of the country, while performance in ES2 (North-East) is comparable to performance in the remaining regions.

In the second part of the table (II), we divide the score gap into three parts: endowments, coefficients and interaction. More formally, we can write: PISA score gap = endowments + coefficients + interaction.

---

[12] Note that it is not possible to compare coefficients directly across regions.

**Table 5.** Blinder-Oaxaca decomposition for Spain

| | ES1 NORTH-WEST | ES2 NORTH-EAST | ES3 MADRID | ES4 CENTRE | ES5 EAST | ES6 SOUTH | ES7 CANARY ISLANDS | |
|---|---|---|---|---|---|---|---|---|
| **Overall** | | | | | | | | |
| Group 1: region | 506.74*** | 495.03*** | 516.14*** | 500.68*** | 499.21*** | 473.24*** | 474.69*** | |
| Group 2: rest of country | 493.34*** | 494.31*** | 490.97*** | 493.52*** | 492.47*** | 501.13*** | 495.40*** | I |
| Difference | 13.41*** | 0.71 | 25.17*** | 7.17* | 6.73* | -27.88*** | -20.71*** | |
| Endowments | -0.29 | 7.01*** | 12.91** | -2.34 | 6.07* | -11.95*** | -11.35** | |
| Coefficients | 12.87*** | -5.16* | 13.50*** | 8.96*** | 3.66 | -14.43*** | -9.67*** | II |
| Interaction | 0.82 | -1.14 | -1.24 | 0.55 | -3.01** | -1.50 | 0.31 | |
| **Endowments** | | | | | | | | |
| Truancy | 0.70*** | 0.47** | -0.42 | 0.44* | -0.52* | 0.09 | 0.17 | |
| Epist. beliefs | 0.21 | -0.37 | 1.82*** | 0.51 | 1.11* | -2.32*** | -0.74 | |
| Learning outside school | -0.23* | 0.24** | -0.07 | -0.69*** | 0.64*** | -0.29* | -0.09 | |
| IB teaching | 0.27 | 0.49* | 0.78** | 0.56** | -1.50*** | 0.34 | 0.17 | |
| TD teaching | -1.97*** | -0.30* | -0.13 | -0.31* | 1.00*** | 0.07 | -0.25 | |
| Science-spec. resources | 0.11 | -0.36 | 0.40 | -0.62 | 1.00 | -0.21 | -0.59 | |
| Min. of science teaching | -0.97* | -0.45 | 3.91*** | -0.04 | -1.19** | -0.13 | -1.30 | |
| SEI | -0.12 | 0.75** | 2.46*** | -0.03 | 0.70* | -2.46*** | -1.26* | |
| Grade repetition | 0.16 | 4.12*** | 0.98 | -2.21* | 4.67*** | -4.77*** | -4.41** | |
| VET | 0.13 | -0.01 | 0.18* | 0.00 | -0.08 | -0.11 | 0.03 | |
| ESCS | 0.46 | 2.35*** | 4.68*** | -0.95 | 1.16 | -3.88*** | -2.88** | |
| Immigrant background | 0.96*** | 0.07 | -1.68*** | 1.02*** | -0.92*** | 1.71*** | -0.19 | |
| **Coefficients** | | | | | | | | |
| Truancy | 0.04 | -0.07 | 0.17 | -0.10 | 0.06 | -0.01 | -0.11 | III |
| Epist. beliefs | 0.02 | 0.26* | -0.08 | 0.07 | -0.16 | -0.03 | 0.11 | |
| Learning outside school | -3.77 | -0.8 | -1.77 | 1.68 | 0.24 | 2.15 | -3.15 | |
| IB teaching | -0.06 | 0.12 | -0.10 | 0.52 | -0.50 | 0.00 | 0.47 | |
| TD teaching | -0.15 | -0.05 | 0.05 | 0.03 | 0.01 | 0.07 | 0.01 | |
| Science-spec. resources | -9.31 | -0.17 | -5.28 | -5.61 | -5.16 | 6.02 | -0.87 | |
| Min. of science teaching | -3.46 | -1.96 | -2.74 | 1.34 | 7.45** | -5.92* | 0.13 | |
| SEI | 13.56* | -3.55 | 5.03 | -3.91 | 4.58 | -10.88 | -20.72*** | |
| Grade repetition | -2.18* | 2.19* | -1.63 | -0.94 | 5.08*** | -3.38* | -3.44** | |
| VET | -0.43* | 0.00 | -0.35 | -0.13 | 0.19 | -0.03 | 0.04 | |
| ESCS | 3.74*** | 1.75* | 0.76 | 1.91** | -0.74 | -0.49 | 0.62 | |
| Immigrant background | -0.29 | -0.17 | 0.08 | 1.10 | -1.32** | 0.34 | 1.19 | |
| Constant | 15.16 | -2.71 | 19.34 | 12.99 | -6.08 | -2.27 | 16.05 | |
| Number of students | 32330 | 32330 | 32330 | 32330 | 32330 | 32330 | 32330 | |

*Note:* Significance level: * p<0.05, ** p<0.01, *** p<0.001.

The first term (endowments) indicates how much of the PISA score gap is due to differences in the characteristics between the rest of the country and the specific region. This could be interpreted as a measure of the mean change of a region's score if it had the same characteristics as the rest of the country (see also Castellano, Longobardi, & Punzo, 2012; Fonseca, Mullen, Zamarro, & Zissimopoulos, 2012; Jann, 2008). The value of -0.29 for ES1 indicates that differences in the endowments (considered together) do not account for any significant share of the score gap. Similarly, endowments do not play a role in explaining performance differences for the ES4 (Centre) region. However, endowments play a significant positive role in the case of ES2 (North-east) and ES3 (Madrid), while they are negatively significant in ES5 (East), ES6 (South) and ES7 (Canary Islands). For example, endowments account for about half of the difference in the Canary Islands (ES7) and Madrid (ES3), while they even make up 90 % of this difference in the East (ES5).

The second term (coefficients) is called the "return" effect and it reflects how much the difference in PISA scores is due to differences in the coefficients on the explanatory variable between the two groups. This could also be interpreted as a measure of how much the region's average student score would change if it had the same average association with performance as in the rest of the country (given the endowments of each region). It can reflect higher or lower effectiveness of using a resource (for example, it may indicate a stronger effect of teaching resources) or socio-economic gradients (for example, it reflects a stronger effect of having a disadvantageous student background). This effect accounts for a significant share of the score gap (around 95%) in the ES1 (North-west) region. This effect also exists in the case of ES3 (Madrid). For ES6 (South) and ES7 (Canary Islands), the coefficients are strongly negatively significant. We also find an intermediate case for ES2, North-east, which is a region with more endowments than the rest of Spain but with lower coefficients relative to the remaining regions. The detailed analysis of coefficients effects provided below allows the interpretation of these results.

The third term (interaction) is the remaining residual of the Blinder-Oaxaca decomposition and shows the multiplication effect of endowments and coefficients. This effect is significant in ES5 region only and explains only small part of the performance gap. Thus, we do not discuss it below.

Finally, we consider the individual variable level (III). While the overall effect of endowments might be insignificant, differences in endowments in individual variables might be significantly associated with differences in science performance. This happens, for instance, when the effect of one variable cancels the effect of another variable (e.g. the positive association of PISA scores with one indicator is balanced by the negative association with another indicator). Thus, it is important to look at individual variables to understand which of them contribute to differences in science performance between regions. We will proceed region by region and we only comment on the results for significant coefficients.

In ES1 (North-west), lower values for TD teaching and for minutes of science teaching and higher values of learning outside school have a negative effect on the score gap. At the same time, the lower values for truancy and immigrant status positively affect the gap[13]. As for the coefficients, we find larger and positive values for both SEI and ESCS, relative to the rest of the country.

As for ES2 (North-east), positive effects on PISA scores are driven by higher values for SEI and ESCS and lower values for grade repetition, truancy and learning outside school and IB teaching. Among the coefficients, significant positive estimates are found for ESCS, grade repetition and epistemological beliefs.

In ES3 (Madrid), higher epistemological beliefs, higher minutes of science teaching, higher SEI, higher average ESCS, lower VET and lower IB teaching positively differentiate this region from others, which is opposed by higher immigrant levels. On the other hand, none of the coefficients is significant.

In ES4 (Centre), high levels for learning outside schools and grade repetition and low values for TD teaching have a negative effect, while low shares of immigrants, low values

[13] By positively (negatively) affecting the score gap we mean that a given variable contributes positively (negatively) to the difference between the average PISA test score of a given region and the one of the comparison group.

for truancy and for IB teaching have a positive effect on the score gap. Among coefficients, only ESCS enters with a positive and significant value.

In ES5 (East), higher values for immigrant status, IB teaching and truancy as well as lower values for minutes of science have a negative effect, while lower levels of learning outside school time, higher values for TD teaching, lower grade repetition and higher SEI have a positive influence on its score gap. Among coefficients, minutes of science and grade repetition have positive and significant values, while immigrant background has a negative value.

In ES6 (South), the negative score gap is associated to lower SEI and higher grade repetition levels, lower epistemological beliefs, lower socio-economic background and higher learning outside school. The smaller share of immigrants is positively characterising this region. Among the coefficients, only minutes of science and grade repetition enter with significant and negative values.

Finally, ES7 (Canary Islands) features higher grade repetition rates, lower ESCS and lower SEI, all of which work to its detriment. No significant values for the coefficients appear.

Overall, the endowments that are most often relevant for within-country differences in our Blinder-Oaxaca decomposition models for Spain are: SEI, grade repetition, truancy, ESCS, immigrant background, learning outside school, IB and TD teaching. Of these, only teaching models are under direct control by the school system, while occupational/educational expectations, grade repetition, learning outside school and truancy are indirectly affected by policies.

**Table 6.** Blinder-Oaxaca decomposition for Italy

| | ITC NORTH-WEST | ITF SOUTH | ITG ISLANDS | ITH NORTH-EAST | ITI CENTRE | |
|---|---|---|---|---|---|---|
| **Overall** | | | | | | |
| Group 1: region | 498.77*** | 459.09*** | 434.72*** | 522.57*** | 481.79*** | |
| Group 2: rest of country | 475.33*** | 486.84*** | 489.97*** | 471.36*** | 480.72*** | I |
| Difference | 23.44*** | -27.75*** | -55.25*** | 51.21*** | 1.07 | |
| Endowments | 2.29 | -9.69 | -8.11 | 11.53 | 3.43 | |
| Coefficients | 23.14*** | -23.05*** | -49.96*** | 42.68*** | -2.32 | II |
| Interaction | -1.99 | 4.99* | 2.83 | -3.00 | -0.03 | |
| **Endowments** | | | | | | |
| Truancy | 1.00 | -2.00*** | -1.64** | 2.87*** | -0.56 | |
| Epist. beliefs | 1.08 | -1.91* | -0.72 | 2.18* | -0.78 | |
| Learning outside school | 0.94** | -1.41*** | -0.79* | 0.82** | -0.55 | |
| IB teaching | 2.48*** | -1.75** | -2.51*** | 0.93 | 1.17 | |
| TD teaching | -1.02* | 0.45 | 0.59 | 0.40 | -0.21 | |
| Science-spec. resources | -0.32 | -2.99 | 1.57 | 1.54 | 0.52 | |
| Min. of science teaching | -0.35 | 0.65 | -0.95 | 0.60 | -0.15 | |
| SEI | -1.14 | 0.50 | -0.08 | 0.22 | 0.72 | |
| Grade repetition | -1.62* | 1.10 | -0.36 | 0.83 | 0.06 | |
| VET | -0.08 | -0.11 | 0.11 | -0.84 | 0.59 | |
| ESCS | 1.59 | -3.17** | -3.86** | 2.28* | 2.78 | |
| Immigrant background | -0.28 | 0.94** | 0.53* | -0.31 | -0.18 | |
| **Coefficients** | | | | | | |
| Truancy | 0.43 | 0.65 | 0.41 | 1.24 | 1.23 | III |
| Epist. beliefs | 0.84 | 0.29 | 0.26 | -0.60 | -0.35 | |
| Learning outside school | -4.39 | 12.28*** | 6.69 | -2.18 | 4.52 | |
| IB teaching | -0.26 | -0.44 | 1.38 | -2.86*** | -0.99 | |
| TD teaching | -0.19 | -0.02 | -0.17 | 0.73 | 0.17 | |
| Science-spec. resources | 2.98 | -20.68 | -1.25 | -1.61 | 14.90 | |
| Min. of science teaching | 3.88 | 3.37 | -10.96*** | -3.26 | 2.08 | |
| SEI | 11.25 | 8.23 | 5.57 | 0.60 | -12.69 | |
| Grade repetition | 0.89 | 0.20 | -0.81 | -0.54 | -0.88 | |
| VET | -4.80 | -5.76 | 0.40 | 13.35* | -11.75 | |
| ESCS | 0.56 | 0.15 | 0.24 | 0.33 | 0.42 | |
| Immigrant background | -1.22* | 1.45 | -0.73 | -2.69*** | -0.01 | |
| Constant | 13.17 | -22.79 | -50.99* | 40.16 | 1.02 | |
| Number of students | 11583 | 11583 | 11583 | 11583 | 11583 | |

*Note:* Significance level: * p<0.05, ** p<0.01, *** p<0.001.

In the case of Italy, we find in the upper part of Table 6 (I) the 'classical' North-South divide: while regions in the North (ITC, ITH) have higher scores than the rest of Italy (499 for the North-west, 523 for the North-east), the students in the South and Islands regions (ITF, ITG) perform lower than the rest of the country (459 in the South, 435 in the Islands). In the same vein, the Centre (ITI) has a score that is not significantly different from the rest of Italy (482 compared to 481). In consequence, the geographical centre of Italy is also an average performing Italian region and there are not any significant coefficients in the following analyses for this region.

The second part of the table (II) shows that the driving factor for the large score gaps is the return on the available educational resources. For example, when considered together, endowments account only for about 10 % of the difference in the North-west (ITC) (notice that overall they are not significant, but single individual endowments do play a role; see III). On the other hand, in Italy returns to coefficients do play an important role, especially in the North-east (ITH) (where they are overall positive) and in the South (ITF) and Islands (ITG) (where they are overall negative).

What variables drive these results? In the North-west (ITC), lower levels of IB teaching and lower levels of learning outside school positively affect the score gap, while negative effects are arising from lower values of TD teaching and higher levels of grade repetition. Among coefficients, only immigrant background enters with a negative and significant value.

In the South (ITF), the score gap is positively affected by lower levels for immigrant background and negatively influenced by higher values for truancy, learning outside school and IB teaching, and by lower values for ESCS and epistemological beliefs. Among coefficients, a large and positive value is found for learning outside school.

In the Islands region (ITG), lower ESCS, higher levels of IB teaching, truancy and learning outside school are negatively associated with the score gap, while a positive effect is coming from lower values for immigrant background. A strong negative association between the minutes of science teaching and scores is also found.

In the North-East (ITH), we find positive associations with the score gap arising from lower levels of truancy, higher ESCS, higher epistemological beliefs and lower learning outside school. Among coefficients, we find a negative association between IB teaching and immigrant background on the one hand and scores on the other hand (but a positive one for VET).

Finally, in the Centre (ITI), as mentioned above, there are no significant results.

Taking all results together, the most relevant endowments revealed by the Blinder-Oaxaca analysis are: learning outside school, truancy, IB teaching and, as expected, the students' socio-economic and immigration background. As for Spain, only IB teaching is under direct influence by education policy, while the effects of the latter on learning outside school and truancy are mediated by students' and households' behaviour.

## 3.3 Discussion of the results

Our descriptive and regression results suggest significant regional inequalities within Italy and Spain. In Spain, the top-performer among NUTS1 regions is the capital area, Madrid, while in Italy it is the North-east. In both MS, the Southern and Island regions perform lower than the other parts of the country. While in Italy the North-South divide is very clear and large, in Spain the regional differences are less pronounced and the geographical pattern is more diverse.

The OLS models show that a range of variables are associated to PISA scores at the NUTS1 level in both MS. In particular, in all regions, teacher-directed instruction and epistemological beliefs are significantly positively associated with students' performance, while a negative association is found for truancy and grade repetition. Nonetheless, other variables are significant, depending on the specific region and each MS. Among them, the most recurrent ones are: students' expected occupational status, minutes of science teaching, inquiry-based teaching, immigrant and socio-economic background and VET attendance.

The Oaxaca models decomposing regional science performance gaps within countries in each MS attenuate the picture by suggesting the role of specific factors that influence within-country score gaps. For example, grade repetition remains a clear factor associated with within-country differences in Spain. On the other hand, while truancy often affects regional disparities in Italy and Spain, in Spain there are other factors such as students' occupational expectations that matter more. The socio-economic and immigrant background also play a large role in many regions. Additionally, lower values for inquiry-based teaching levels can be helpful in positively distinguishing a region from the rest of the country.

These results are also mostly in line with the literature. Socio-economic background is usually considered a strong predictor of a student's score (Agasisti & Cordero-Ferrera, 2013; Willms, 2006; Wößmann, 2007), which we also find in our analyses. There has been some discussion on the impact of immigrant status on PISA scores in Spain (Ciccone & Garcia-Fontes, 2009; Cordero Ferrera, Crespo Cebada, Pedraja Chaparro, & Santín González, 2011; Ferrer, Valiente, & Castel, 2010; Méndez, Zamarro, Clavel, & Hitt, 2015). Our results indicate that students with an immigrant background are often associated to negative score gaps within the country. Similarly, grade repetition is a significant factor leading to within-country score differences in Spain. Thus, our results in this area confirm the finding of Cordero Ferrera, Crespo Cebada, Pedraja Chaparro, & Santín González (2011).

Interestingly, in both MS learning outside school time and inquiry-based teaching are (at least in some cases) relevant factors for score differences within MS. Learning outside school time is a rather ambiguous variable, as it can be related to inefficiencies of schools, but also to individual problems students experience with learning; it can also reflect that students with larger family resources take private lessons more often. Thus, more specific information would be needed on the way this additional learning time is spent, but this is not analysed in this paper due to data limitations. Inquiry-based teaching has also been shown by the OECD to be negatively related to science scores (OECD, 2016d), confirming our results. On the other hand, (occupational) expectations about the future do matter in Spain, while in Italy we could not detect their effect. Another contribution of our study is our self-computed truancy variable, which is shown to be relevant in some regions.

Other researchers have indicated the negative effects of school tracking in Italy (Checchi, 2004; Fornari & Giancola, 2011). Indeed, we also find that students from vocational schools have lower scores than pupils on general tracks in PISA 2015. However, our decomposition analysis does not indicate that this is a relevant factor for within-country differences. Similarly, various researchers (Agasisti & Cordero-Ferrera, 2013; Bratti, Checchi, & Filippin, 2007; Dolton & Marcenaro-Gutierrez, 2011; Donato & Ferrer-Esteban, 2012; Fuentes, 2009; Hanushek & Luque, 2003) discuss the relevance of the allocation of resources and educational spending in explaining regional or cross-country differences, and sometimes find positive, sometimes insignificant effects. In general, while our results indicate that endowments often play a significant role in the Spanish regions (and a smaller one in Italy), we do not find evidence that science-specific resources are relevant for within-country differences in any of the two MS. Still, it is possible that other resources not included in our analyses may matter.

Given the low number of regions in our study, more advanced multilevel analyses (as an alternative to OLS regressions) could not be run, and this prevented us from considering several levels (e.g., including also the regional or the school level) in our analyses. In the same vein, studies on earlier PISA rounds were able to exploit regional differences at the NUTS2 level in Italy (e.g., Agasisti & Cordero-Ferrera, 2013; OECD, 2014; Sibiano & Agasisti, 2013). Such a detailed analysis would have allowed us to point out more specifically which parts of the macro-regions are driving the results, and to investigate possible inequalities within these larger areas. However, using PISA 2015 we have to limit ourselves to the macro-regional level to employ the same unit of analysis in both MS.

# 4 Policy messages

This analysis shows significant regional differences within Spain and Italy. Therefore, the results indicate that, even if the national scores of a MS are not outstanding, it is possible that some regions can obtain excellent results. Thus, national averages do not give a detailed enough picture of the actual performance attained within a MS. In consequence, national education policies which are only based on national comparative data run the risk of not achieving their intended goals by not taking into account the specificities of the regions within a MS.

Moreover, comparing results of MS and relating them to national policies might be misleading as this brief suggests that some educational practices and traditions vary across regions within MS and that some of these measurable characteristics are related to regional differences in students' performance. The brief reveals a more complex structure for the implementation of educational policies that should recognize distinct features of regions. Indeed, the implementation and consequences of education policies at the regional level may not always give the same outcomes, as the PISA results indicate.

In consequence, employing country averages from international large-scale assessments like PISA may not be useful for a country like Italy, where results vary substantially across macro-regions. For example, low truancy levels, high epistemological beliefs and low learning outside school time provide advantages to the North-east, while opposite values in these variables are associated to the lagging behind of the South. Still, the way in which available resources are used and spent also matters in determining regional PISA score differences in this MS.

In Spain, the picture is more diverse than in Italy. Endowments, considered together, play a bigger role in explaining the score gap (i.e., North-east (ES2), Madrid (ES3), South (ES6) and Canary Islands (ES7)). Expectations (about future occupations) are important, as expectations are contributing to the score gap in the rest of the country. The significant role of students' (occupational) expectations at the regional level may be indicative of a variety of underlying causes, such as the motivations of a student population in a given region, or the characteristics of the given economic sectors within a region and what they are perceived to offer to students. Raising expectations is certainly not easy, but showing students the prospects of certain careers and informing them about the possibilities offered by higher level occupations could increase the willingness of students to study and learn for their future integration in the labour market.

At the same time, having repeated a grade has often a significant effect on the regional PISA score gaps. In particular, it appears that grade repetition practices are different among the various Spanish regions, and their effect appears to vary substantially across regions, so that policy makers may consider the efficiency or lack thereof of these practices more closely.

The contribution of immigrant students to the PISA score gap depends on the region. Immigrants have lower scores in all regions, but the score gap in relation to that of native students can be quite different in each region. At the same time, the integration of immigrants needs to be improved, in particular in the North-east region in Italy and the East of Spain.

Furthermore, implementing policies that limit truancy (skipping classes, etc.) appears to be a potentially effective way to improve regional performance in some regions. In addition, it may be relevant to reconsider the concepts and practices of inquiry-based teaching, as we have found that it is associated with lower regional performance. Indeed, this is a new finding, as the OECD mentions this only for the country level. In our case, it is consistent across regions.

Finally, policies also need to be set in a wider framework addressing e.g., social, financial and cultural policies (see also European Commission, 2013). There does not appear to be a simple and clear-cut 'magic' policy to raise educational achievement within a region and to overcome regional differences within a country. Instead, education policies may need to be geared to the existing local and regional necessities and availabilities. In fact, they may include all relevant stakeholders in addition to educational institutions – e.g. considering the needs of parents, the resources of local governments, the skills needs and job opportunities of local businesses – in an overarching framework for the provision and improvement of local education. Schools are certainly the key institution in the provision of education and skills, but they are interdependent with all the other (local) actors. Therefore, it may be useful to understand in-depth the vested interests of local and regional actors and of the incentives that students have for learning, and the obstacles that need to be removed for increasing their performance, in each region.

# 5 Future research avenues

While the socio-economic background (ESCS) is well-known to impact science scores, epistemological beliefs are included in PISA 2015 for the first time. It appears that students' perception and understanding of science is quite important. In the same vein, inquiry-based teaching has been shown to have a significant (negative) association with regional PISA results. This brief finds that these factors differ across regions and that they are associated with regional differences in student performance. More research is needed to understand these complex relationships, however, especially the interaction between national and regional policies in both MS. Similarly, more research on the expectations of students and on truancy would be useful to create better policies and allow all students to achieve their full potential.

This study has important limitations, which should be addressed in future research and in future PISA rounds. Regional information on PISA scores is only available for large macro-areas in Italy in 2015, while more detailed data for NUTS2 regions were included in previous rounds. Indeed, the analysis would have benefitted from regional data at the NUTS2 level if it had been available for both MS. Given the fact that we already find a substantial variance in PISA scores by considering the more aggregated regional level in both MS, our results indicate the usefulness and promises of regional analyses using PISA and that even higher differences may appear at lower regional levels. Similarly, the Blinder-Oaxaca decomposition models would have potentially been able to generate more significant values with more observations, in particular in Italy. However, to conduct a more in-depth study of regional inequalities, more regional data would need to be collected in future PISA rounds in these (and other) MS.

## References


Agasisti, T., & Cordero-Ferrera, J. M. (2013). Educational disparities across regions: A multilevel analysis for Italy and Spain. *Journal of Policy Modeling*, *35*(6), 1079–1102.

Baten, J., & Hippe, R. (2017). Geography, Land Inequality and Regional Numeracy in Europe in Historical Perspective. *Journal of Economic Growth*, 23(1): 79-109.

Benhenda, A., & Grenet, J. (2015). How much does grade repetition in French primary and secondary schools cost? IPP Note, 17/2015, The Institut des Politiques Publiques, Paris.

Blinder, A. S. (1973). Wage discimination: Reduced form and structural estimates. *Journal of Human Resources*, *8*, 436–455.

Bratti, M., Checchi, D., & Filippin, A. (2007). Geographical differences in Italian students' mathematical competencies: evidence from PISA 2003. *Giornale Degli Economisti E Annali Di Economia*, 299–333.

Cappelli, G. (2015). Escaping from a human capital trap? Italy's regions and the move to centralized primary schooling, 1861–1936. *European Review of Economic History*, *20*(1), 46–65.

Cappelli, G. (2016). One size that didn't fit all? Electoral franchise, fiscal capacity and the rise of mass schooling across Italy's provinces, 1870–1911. *Cliometrica*, *10*(3), 311–343.

Castellano, R., Longobardi, S., & Punzo, G. (2012). Castellano, R., Longobardi, S., & Punzo, G. (2012). Do Italian students perform worse than their OECD fellows? A decomposition analysis of educational gaps. *Italian Journal of Applied Statistics*, *22*(2), 99–127.

Checchi, D. (2004). Da dove vengono le competenze scolastiche? *Stato E Mercato*, *24*(3), 413–454.

Ciccone, A., & Garcia-Fontes, W. (2009). The quality of the Catalan and Spanish education systems: A perspective from PISA. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1513214

Cordero Ferrera, J., Crespo Cebada, E., Pedraja Chaparro, F., & Santín González, D. (2011). Exploring educational efficiency divergences across Spanish regions in PISA 2006. *Revista de Economía Aplicada*, *19*(57). Retrieved from http://www.redalyc.org/html/969/96922243005/

Diebolt, C., & Hippe, R. (2018). Remoteness equals backwardness? Human capital and market access in the European regions: insights from the long run. *Education Economics,* 26 (3): 285-304.

Dolton, P., & Marcenaro-Gutierrez, O. D. (2011). If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance. *Economic Policy*, *26*(65), 5–55.

Donato, L., & Ferrer-Esteban, G. (2012). Desigualdades territoriales en España e Italia: nuevas evidencias a partir de la evaluación PISA-2009. *Revista Española de Educación Comparada*, (19), 105–138.

European Commission. (2013). Reducing early school leaving: Key messages and policy support. Final Report of the Thematic Working Group on Early School Leaving. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/policy/strategic-framework/doc/esl-group-report_en.pdf

Eurostat. (2016). National structures (EU). Retrieved from http://ec.europa.eu/eurostat/web/nuts/national-structures-eu

Eurostat. (2017). Principles and characteristics,. Retrieved from http://ec.europa.eu/eurostat/web/nuts/principles-and-characteristics

Ferrer, F., Valiente, Ó., & Castel, J. L. (2010). Los resultados PISA-2006 desde la perspectiva de las desigualdades educativas: la comparación entre Comunidades Autónomas en España. *Revista Española de Pedagogía*, 23–47.

Financial Times. (2016). Widening regional disparities in Europe. Retrieved from https://www.ft.com/content/539593ff-68ce-34a3-a672-20398c832c19

Fonseca, R., Mullen, K. J., Zamarro, G., & Zissimopoulos, J. (2012). What explains the gender gap in financial literacy? The role of household decision making. *Journal of Consumer Affairs*, *46*(1), 90–106.

Fornari, R., & Giancola, O. (2011). Policies for decentralization, school autonomy and educational inequalities among the Italian regions. Empirical evidence from Pisa 2006. *Italian Journal of Sociology of Education*, *3*(2).

Fuentes, A. (2009). Raising education outcomes in Spain. *OECD Economic Department Working Papers*, (666), 0_1.

Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, *21*(1), 1–56.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2013). Human Capital and Regional Development. *Quarterly Journal of Economics*, *128*(1), 105–164.

Hanushek, E. A., & Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of Education Review*, *22*(5), 481–502.

Hanushek, E. A., & Wößmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. MIT Press.

Hippe, R., Araújo, L., & Dinis da Costa, P. (2016). Equity in Education in Europe. Retrieved from Luxembourg (Luxembourg): Publications Office of the European Union; EUR 28285 EN; doi:10.2791/255948

Ikeda, M., & García, E. (2014). Grade repetition. *OECD Journal: Economic Studies*, *2013*(1), 269–315.

Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *Stata Journal*, *8*(4), 453–479.

Kolenikov, S., & Ángeles, G. (2004). *The use of discrete data in principal component analysis with applications to socio-economic indices. CPC*. MEASURE Working paper no. WP-04-85.

Matteucci, M., & Mignani, S. (2014). Exploring regional differences in the reading competencies of Italian students. *Evaluation Review*, *38*(3), 251–290.

Méndez, I., Zamarro, G., Clavel, J. G., & Hitt, C. (2015). Non-cognitive Abilities and Spanish Regional Differences in Student Performance in PISA 2009. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2652322

Oaxaca, R. (1973). Male–female wage differentials in urban labor markets. *International Economic Review*, *14*, 693–709.

OECD. (2010). OECD Reviews of Migrant Education - Closing the Gap for Immigrant Students: Policies, Practice and Performance. Retrieved from http://www.oecd.org/edu/school/oecdreviewsofmigranteducation-closingthegapforimmigrantstudentspoliciespracticeandperformance.htm

OECD. (2014). PISA 2012 Results: What Students Know and Can Do. Student peformance in mathematics, reading and science. OECD Publishing.

OECD. (2016a). *OECD Regions at a Glance 2016.* Paris: OECD Publishing. Retrieved from http://dx.doi.org/10.1787/reg_glance-2016-en

OECD. (2016b). PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy. OECD Publishing, Paris. Retrieved from http://dx.doi.org/10.1787/9789264255425-en

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

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

OECD. (2016e). PISA country note. Japan. Retrieved from https://www.oecd.org/pisa/PISA-2015-Japan.pdf

OECD. (2017a). PISA 2015 database. Codebokes for the main files. Retrieved from http://www.oecd.org/pisa/data/2015database/Codebook_CMB.xlsx

OECD. (2017b). PISA 2015. Technical Report. Chapter 16: Procedures and Construct Validation of Context Questionnaire Data. Retrieved from http://www.oecd.org/pisa/data/2015-technical-report/

Royston, P. (2009). Multiple imputation of missing values: further update of ice, with an emphasis on categorical variables. *Stata Journal*, *9*(3), 466.

Ruhose, J., & Schwerdt, G. (2016). Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries. *Economics of Education Review*, *52*, 134–154.

Sibiano, P., & Agasisti, T. (2013). Efficiency and heterogeneity of public spending in education among Italian regions. *Journal of Public Affairs*, *13*(1), 12–22.

Willms, J. D. (2006). Learning divides: ten policy questions about the performance and equity of schools and school systems. Montreal: UNESCO Institute of Statistics.

Wößmann, L. (2007). Fundamental determinants of school efficiency and equity: German states as a microcosm for OECD countries. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=986555

## List of abbreviations and definitions

### A. EU country abbreviations

| Code | Country name |
|---|---|
| ES | Spain |
| IT | Italy |

### B. Other abbreviations

| Abbreviation | Description |
|---|---|
| ESCS | Economic, Social and Cultural Status |
| EU | European Union |
| GDP | Gross Domestic Product |
| ILSA | International Large Scale Assessments |
| IRT | Items Response Theory |
| ISCED | International Standard Classification of Education |
| NUTS | Nomenclature for Territorial Statistics |
| MS | Member States |
| OECD | Organisation for Economic Co-operation and Development |
| PISA | Programme for International Student Assessment |
| S.E. | Standard Errors |

## List of figures

## List of tables

# Appendix

## Annex 1. Details on the regional methodology

*NUTS classification*

In this brief we use the NUTS 2013 classification. NUTS distinguishes various levels that go from NUTS0 (the country level) to NUTS3 (e.g., *provincias* in Spain). Every higher level from NUTS0 to NUTS2 is composed of various regions of the lower level. The NUTS classification of regions follows several criteria: First, NUTS distinguishes regions according to their population. The table below shows the specific thresholds:

| Level | Minimum | Maximum |
|---|---|---|
| NUTS 1 | 3 000 000 | 7 000 000 |
| NUTS 2 | 800 000 | 3 000 000 |
| NUTS 3 | 150 000 | 150 000 |

*Source:* Eurostat (2017).

In practice, this means that a country like Luxembourg is a NUTS3 region. However, as all countries have to have NUTS0 to NUTS3 levels, it is also a NUTS2, NUTS1 and NUTS0 region, even if the population size does not correspond to the criteria shown in the table above.

The second principle relates to the fact that the NUTS classification does not 'invent' new regions but it prefers to take those which are provided by the administrative classification schemes of MS. This makes the classification more practical and relevant to policy makers.

Finally, NUTS codes are regularly updated to be corresponding to the administrative changes that MS implement.

The resulting NUTS regional methodology can be summarised as follows for the case of Italy and Spain (see the table below):

| MS | NUTS 1 | | NUTS 2 | | NUTS 3 | |
|---|---|---|---|---|---|---|
| ES | Agrupacion de comunidades Autonomas | 7 | Comunidades y ciudades Autonomas | 19 | Provincias + islas + Ceuta, Melilla | 59 |
| IT | Gruppi di regioni | 5 | Regioni | 21 | Provincie | 110 |
| *EU-28* | | *98* | | *276* | | *1343* |

*Source:* Eurostat (2016).

*Country-specific details on the regions in PISA 2015*

ES: We have obtained information on additional regional data availability by Carmen Tovar Sánchez.

IT: We have received help for obtaining NUTS1 regions by Carlo Di Chiacchio. Regional data were retrieved from the ISTAT database.

**Annex 2. Details on the included variables**

In the following we provide additional information the variables included in this study, directly taken from and as presented by OECD sources (2017a, 2017b), except for truancy which we constructed ourselves (see main text for more details).

*Truancy*

An index which we constructed using the responses to the test questions below:

| Item | In the last two full weeks of school, how often did the following things occur? |
|---|---|
| ST062Q01TA | In the last two full weeks of school, how often: I <skipped> a whole school day |
| ST062Q02TA | In the last two full weeks of school, how often: I <skipped> some classes |
| ST062Q03TA | In the last two full weeks of school, how often: I arrived late for school |

*Epistemological beliefs*

Epistemological beliefs about science were measured with a new question about students' views on scientific approaches (ST131). Students answered on a four-point Likert scale with the answering categories "strongly agree", "agree", "disagree", and "strongly disagree". The derived variable EPIST was scaled using the IRT scaling model described above. [The table below] shows the item wording, international item parameters and item fit for EPIST.

| Item | How much do you disagree or agree with the statements below? |
|---|---|
| ST131Q01NA | A good way to know if something is true is to do an experiment. |
| ST131Q03NA | Ideas in <broad science> sometimes change. |
| ST131Q04NA | Good answers are based on evidence from many different experiments. |
| ST131Q06NA | It is good to try experiments more than once to make sure of your findings. |
| ST131Q08NA | Sometimes <broad science> scientists change their minds about what is true in science. |
| ST131Q11NA | The ideas in <broad science> science books sometimes change. |

*Immigration background*

The PISA database contains three country-specific variables relating to the students' country of birth, their mother and father (COBN_S, COBN_M, and COBN_F). The items ST019Q01TA, ST019Q01TB and ST019Q01TC were recoded into the following categories: (1) country of birth is the same as country of assessment and (2) other. The index of immigrant background (IMMIG) was calculated from these variables with the following categories: native students (those students who had at least one parent born in the country), (2) second generation students (those born in the country of assessment but whose parent(s) were born in another country) and (3) first-generation students (those students born outside the country of assessment and whose parents were also born in another country). Students with missing responses for either the student or for both parents were assigned missing values for this variable.

*Hours of learning outside school*

Students were asked in a slider-format question how much time they spent studying in addition to their required school schedule (ST071). The index OUTHOURS was computed by summing the time spent studying for different school subjects.

*Student age*

The age of a student (*AGE*) was calculated as the difference between the year and month of the testing and the year and month of a student's birth. Data on student's age were obtained from both the questionnaire (ST003) and the student tracking forms. If the month of testing was not known for a particular student, the median month for that country was used in the calculation. The formula for computing *AGE* was

$AGE = (100 + Ty - Sy) + (Tm - Sm)/12$

where Ty and Sy are the year of the test and the year of the students' birth, respectively in two-digit format (for example "06" or "92"), and Tm and Sm are the month of the test and month of the students' birth, respectively. The result is rounded to two decimal places.

*Expected occupational status*

As in previous cycles of PISA, students were asked to report their expected occupation at age 30 and a description of this job. The responses were coded to four-digit ISCO codes […] and then mapped to the ISEI index […]. Recoding of ISCO codes into ISEI index results in scores for the students' expected occupational status (BSMJ), where higher scores of ISEI indicate higher levels of expected occupational status.

*Early childhood education and care*

Questions ST125 and ST126 measure the starting age in ISCED 1 and ISCED 0. A difference score of the two thus indicates the number of years a student spent in early childhood education and care. This indicator is called DURECEC.

| | |
|---|---|
| ST125Q01NA | How old were you when you started <ISCED 0>? Years |
| ST126Q01TA | How old were you when you started <ISCED 1>? Years |

*Grade repetition*

The grade repetition variable (REPEAT) was computed by recoding variables ST127Q01TA, ST127Q02TA, and ST127Q03TA. REPEAT took the value of "1" if the student had repeated a grade in at least one ISCED level and the value of "0" if "no, never" was chosen at least once, given that none of the repeated grade categories were chosen. The index is assigned a missing value if none of the three categories were ticked in any levels.

| | |
|---|---|
| ST127Q01TA | Have you ever repeated a <grade>? At <ISCED 1> |
| ST127Q02TA | Have you ever repeated a <grade>? At <ISCED 2> |
| ST127Q03TA | Have you ever repeated a <grade>? At <ISCED 3> |

*Minutes of science teaching (SMINS)*

Learning time in test language (LMINS) was computed by multiplying the number of minutes on average in the test language class by number of test language class periods per week (ST061 and ST059). Comparable indices were computed for mathematics (MMINS) and science (SMINS).

*Instruction-based teaching*

For ST098, students responded on a four-point Likert scale with the categories "in all lessons", "in most lessons", "in some lessons", "never or hardly ever". Therefore, the ST098-items were reverse coded so that higher WLEs and higher difficulty correspond to higher levels inquiry-based science teaching and learning practices. [The table below] shows the item wording, international item parameters and item fit for IBTEACH.

| Item | When learning <school science> topics at school, how often do the following activities occur? |
|---|---|
| ST098Q01TA | Students are given opportunities to explain their ideas. |
| ST098Q02TA | Students spend time in the laboratory doing practical experiments. |
| ST098Q03NA | Students are required to argue about science questions. |
| ST098Q05TA | Students are asked to draw conclusions from an experiment they have conducted. |
| ST098Q06TA | The teacher explains how a <school science> idea can be applied to a number of different phenomena (e.g. the movement of objects, substances with similar properties). |
| ST098Q07TA | Students are allowed to design their own experiments. |
| ST098Q08NA | There is a class debate about investigations. |
| ST098Q09TA | The teacher clearly explains the relevance of <broad science> concepts to our lives. |

*Teacher-directed teaching*

For ST103, students responded on a four-point Likert scale with the categories "never or almost never", "some lessons", "many lessons", and "every lesson or almost every lesson". [The table below] shows the item wording, international item parameters and item fit for TDTEACH.

| Item | How often do these things happen in your lessons for this <school science> course? |
|---|---|
| ST103Q01NA | The teacher explains scientific ideas. |
| ST103Q03NA | A whole class discussion takes place with the teacher. |
| ST103Q08NA | The teacher discusses our questions. |
| ST103Q11NA | The teacher demonstrates an idea. |

*Science-specific resources*

A new index was built in 2015 to reflect the schools' science-specific resources (SCIERES). It was constructed by summing up the principals' answers to SC059 (yes/no question)

| | |
|---|---|
| SC059Q01NA | Compared to other departments, our school's <school science department> is well equipped. |
| SC059Q02NA | If we ever have some extra funding, a big share goes into improvement of our <school science> teaching. |
| SC059Q03NA | <School science> teachers are among our best educated staff members. |
| SC059Q04NA | Compared to similar schools, we have a well equipped laboratory. |
| SC059Q05NA | The material for hands-on activities in <school science> is in good shape. |
| SC059Q06NA | We have enough laboratory material that all courses can regularly use it. |
| SC059Q07NA | We have extra laboratory staff that helps support <school science> teaching. |
| SC059Q08NA | Our school spends extra money on up-to-date <school science> equipment. |

*Vocational education (VET)*
Programme orientation (ISCEDO) indicates whether the programme's curricular content was general, pre-vocational or vocational.

## Annex 3. Descriptive statistics and distribution of other variables

**Table 7.** Descriptive statistics for student-level variables in the original dataset

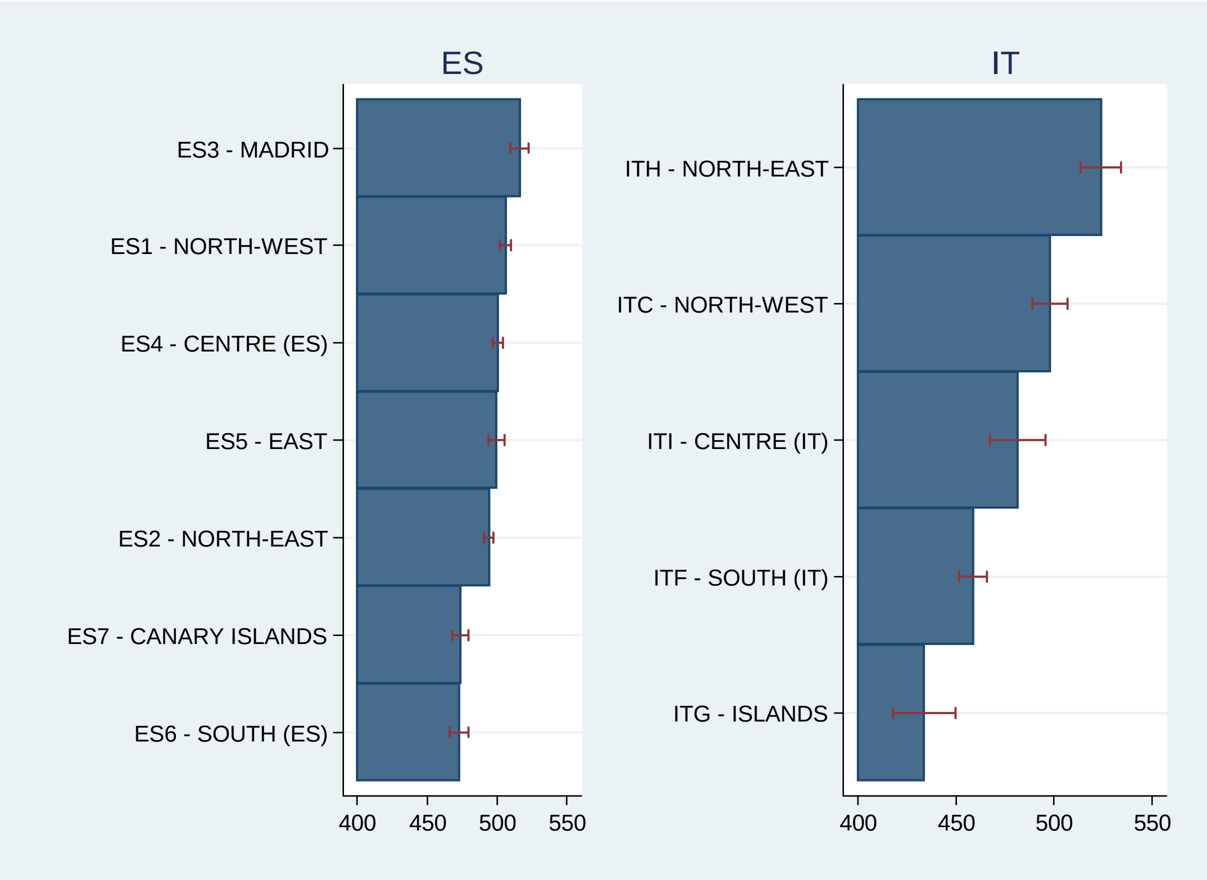
| | Spain | | | Italy | | |
|---|---|---|---|---|---|---|
| | N | Mean | SD | N | Mean | SD |
| Science performance | 32330 | 494.13 | 88.79 | 11583 | 480.70 | 91.77 |
| Truancy | 31549 | -0.08 | 0.99 | 11109 | 0.28 | 1.11 |
| Epistemological beliefs | 30393 | 0.10 | 1.00 | 10804 | -0.10 | 0.88 |
| PISA index of socio-economic status | 32035 | -0.52 | 1.18 | 11330 | -0.07 | 0.95 |
| Student with immigrant background | 31640 | 0.12 | 0.32 | 11232 | 0.08 | 0.27 |
| Hours of learning outside school | 30413 | 18.15 | 12.41 | 10075 | 21.18 | 13.35 |
| Minutes of science teaching | 31257 | 193.07 | 125.53 | 11028 | 155.12 | 121.91 |
| Expected occupational status | 28348 | 60.20 | 16.49 | 9384 | 57.53 | 18.27 |
| Student repeating a grade | 32130 | 0.32 | 0.47 | 11330 | 0.15 | 0.36 |
| Student in vocational education | 32330 | 0.01 | 0.10 | 11583 | 0.50 | 0.50 |
| Inquiry-based teaching | 26551 | -0.27 | 0.96 | 10337 | -0.20 | 0.92 |
| Teacher-directed teaching | 26378 | 0.05 | 0.90 | 10310 | -0.15 | 0.84 |
| Science-specific resources | 30909 | 4.43 | 1.95 | 8219 | 4.71 | 1.77 |

Note that PISA 2015 data contain information for all students about their science performance. However, additional background information from students and school principals have gaps due to missing responses. Data imputation techniques are necessary to address this issue and they can be applied as the amount of missing data per variable is relatively small. In consequence, sample sizes for the original data vary, while for the imputed dataset there is no missing data so full sample size is available. It is worth noting that while there is substantial variation across students in these variables, descriptive statistics for the original and imputed datasets are very close. In our case we used multiple imputation with chained equations that can deal with different data types (continuous, ordered and nominal) (see Royston, 2009).

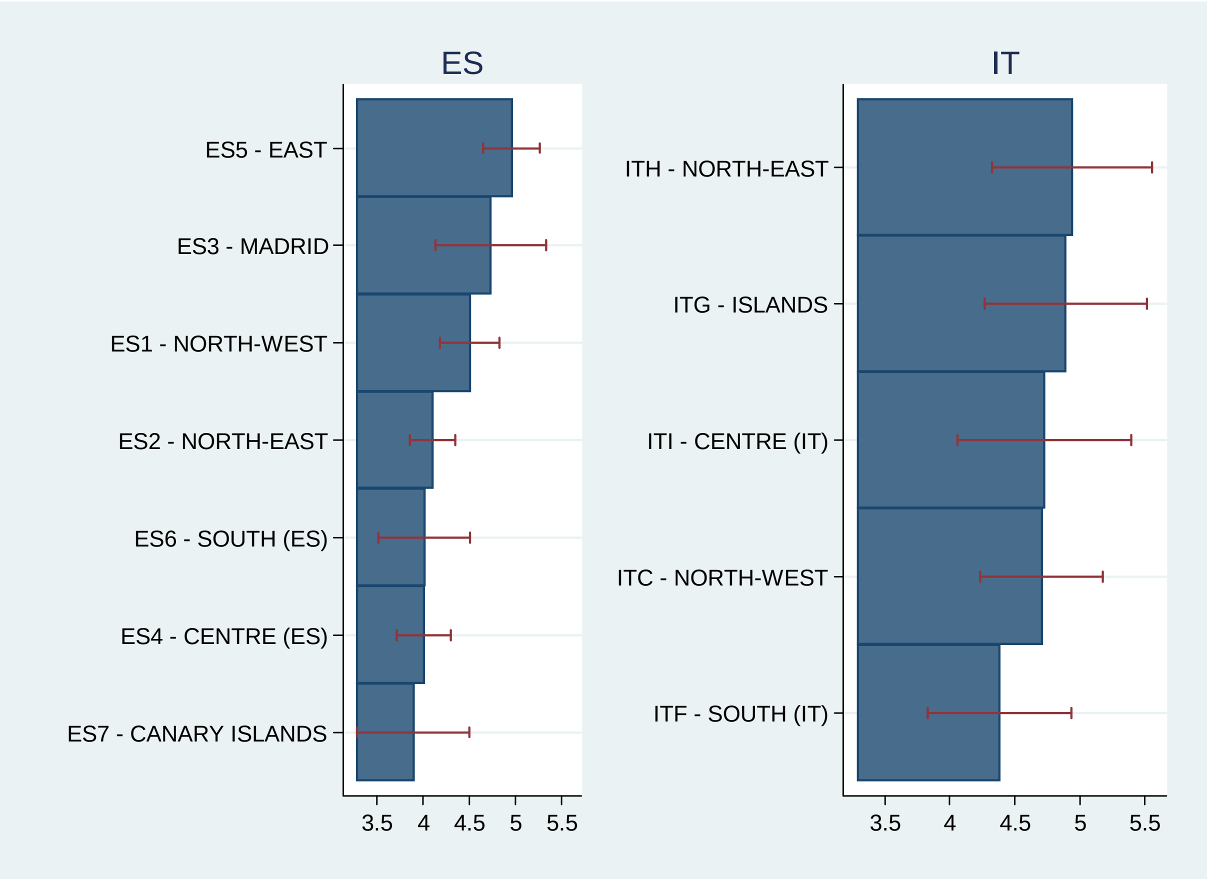
After having presented graphically the regional distribution of some of the most important variables in the main text, we show the distribution of other variables which we use at the regional level in our models. The table below provides the description of the variables shown in the graphs.

| Abbreviation | Variable description |
|---|---|
| science | Science scores |
| scieres | Science-specific resources |
| smins | Minutes of science teaching |
| tdteach | Teacher-directed science instruction |
| voco | Vocational education |

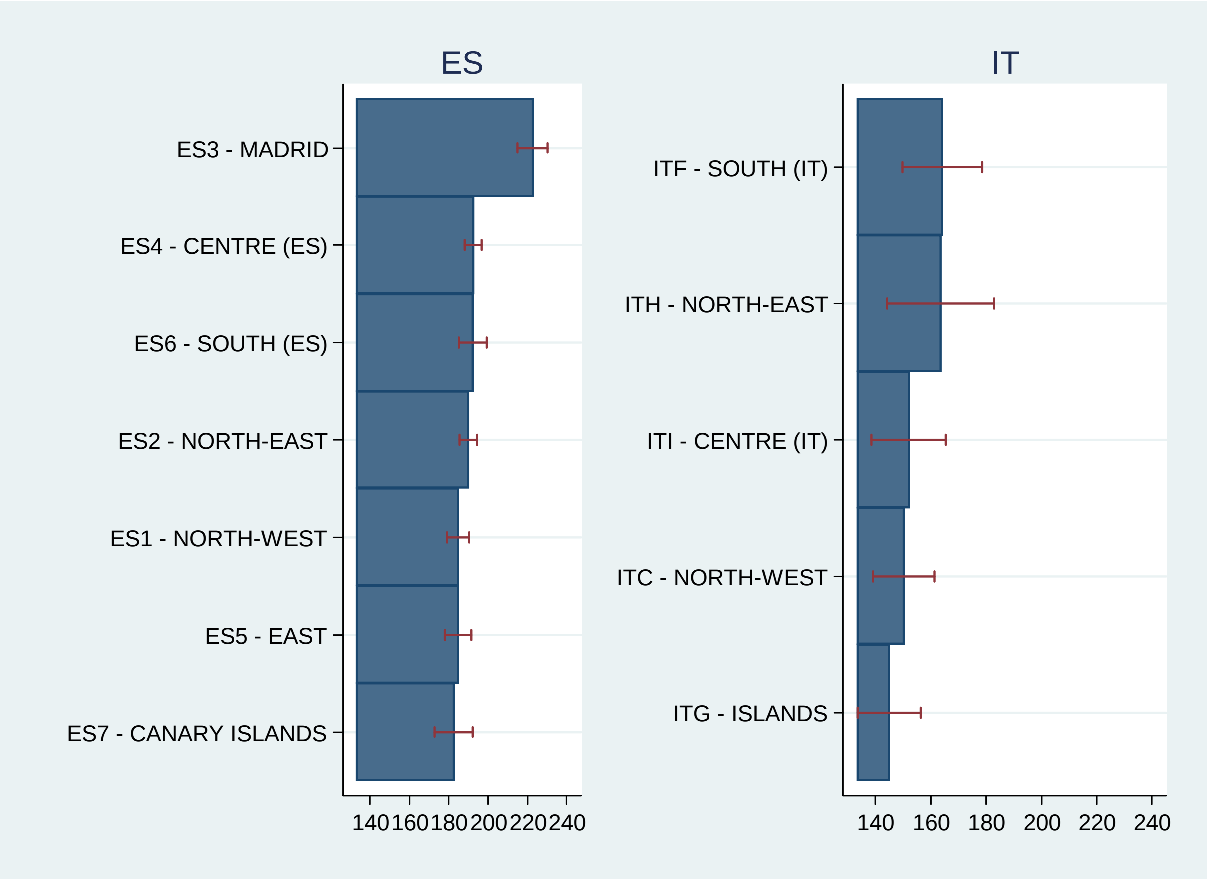
Science scores



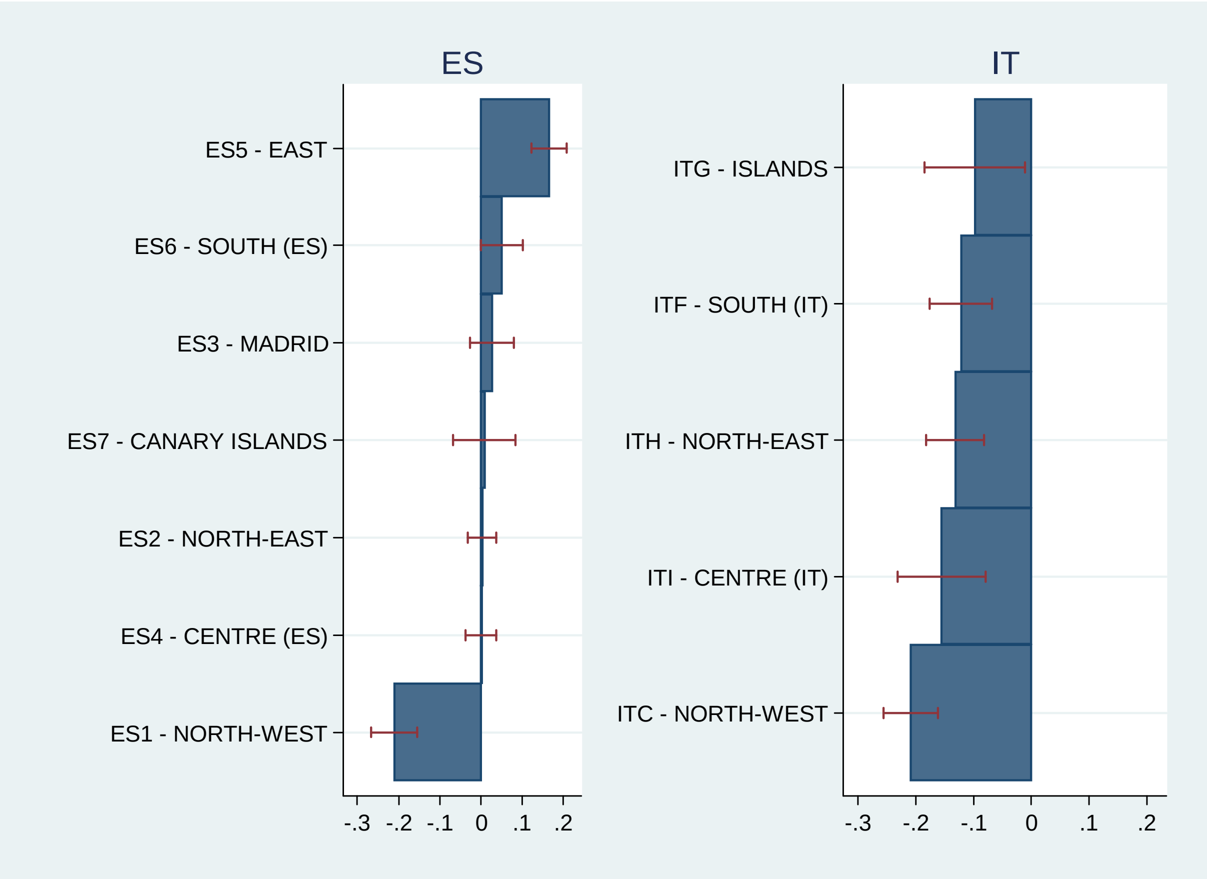
Science-specific resources



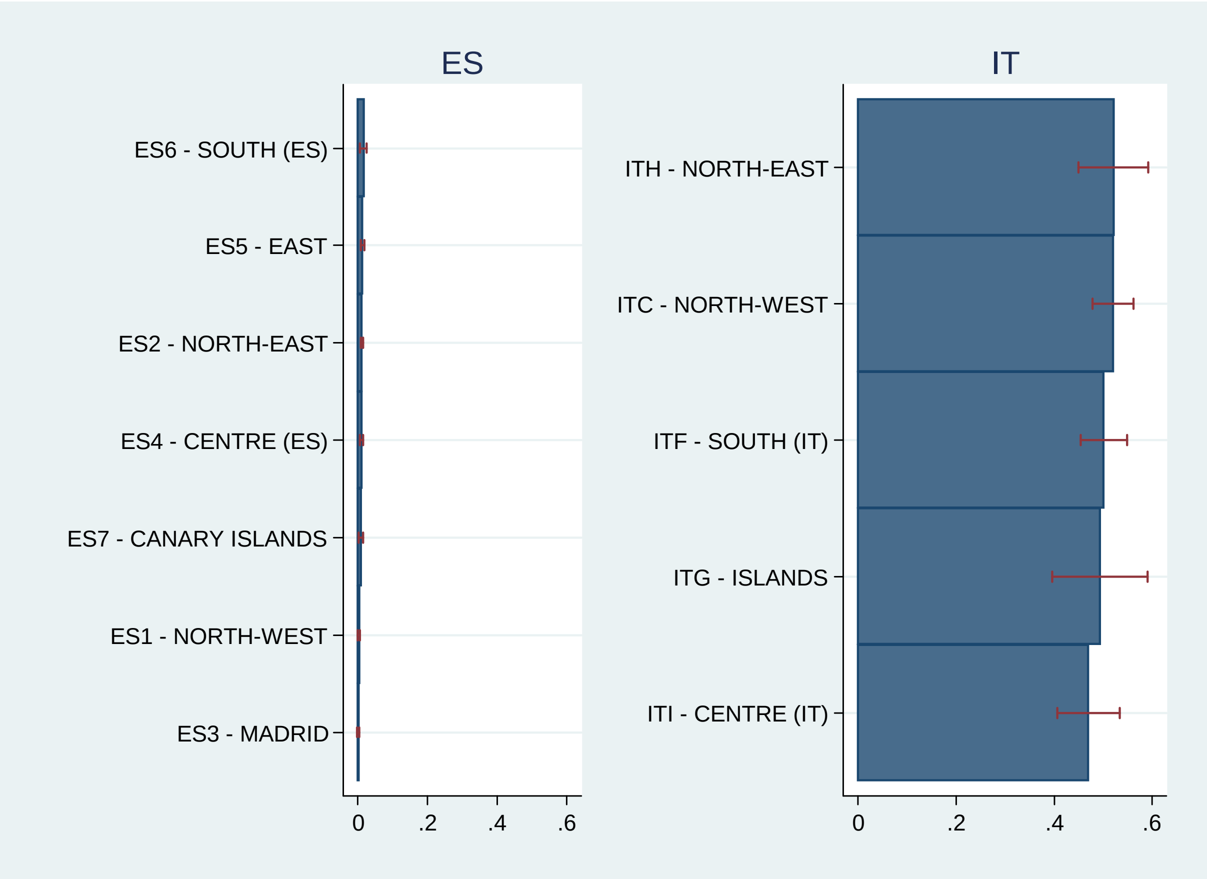
Minutes of science teaching



Teacher-directed science instruction



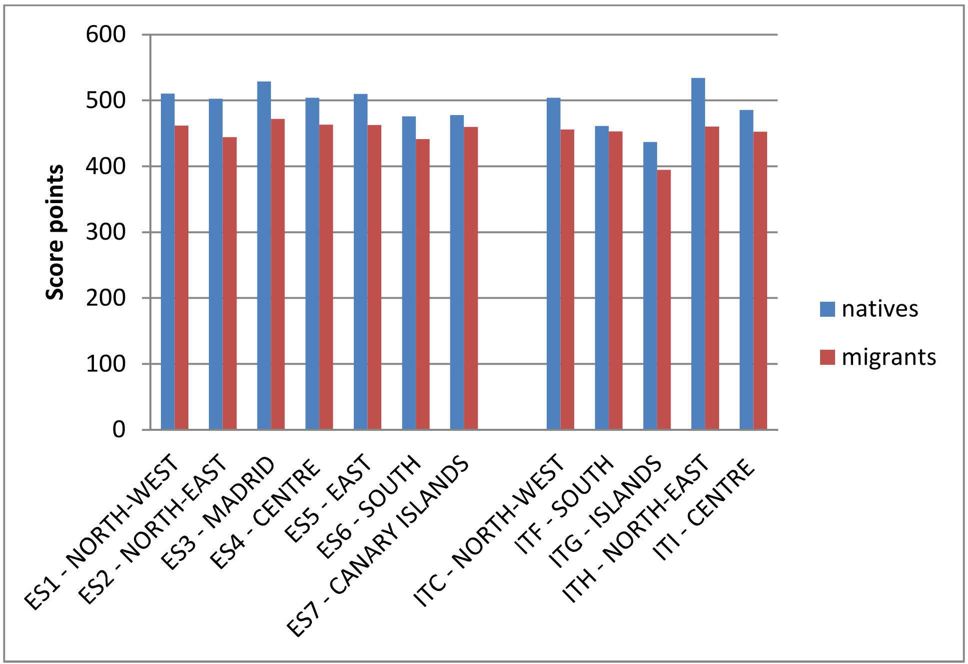
Vocational education

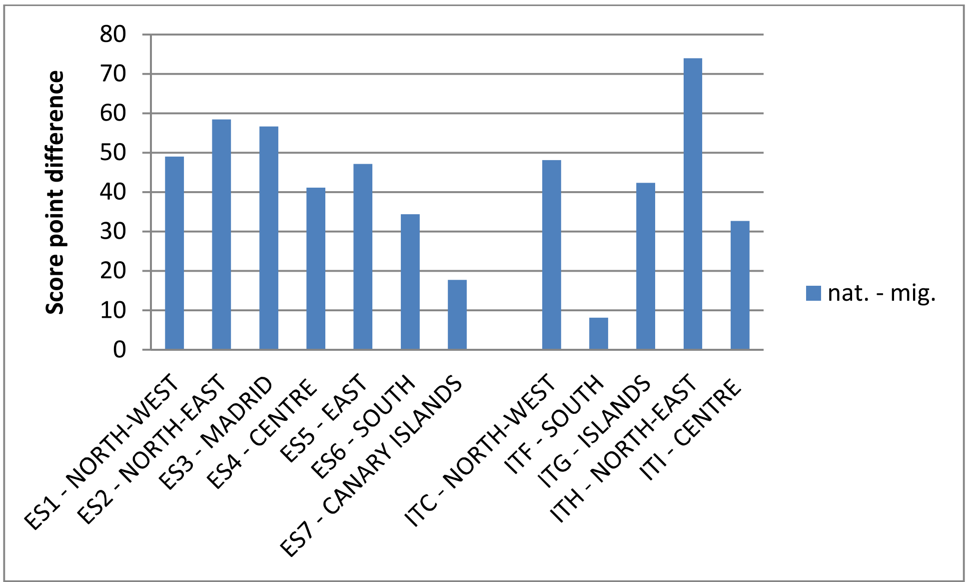


**Annex 4. Further score point differences among categories of variables**

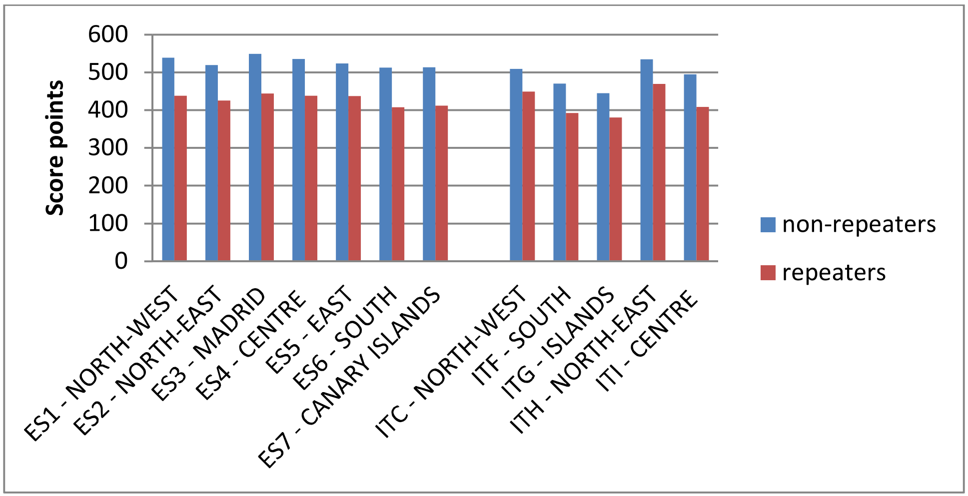
We present in the following a number of graphs that show PISA scores for the different groups of a variable and PISA score differences among the two groups of a variable.

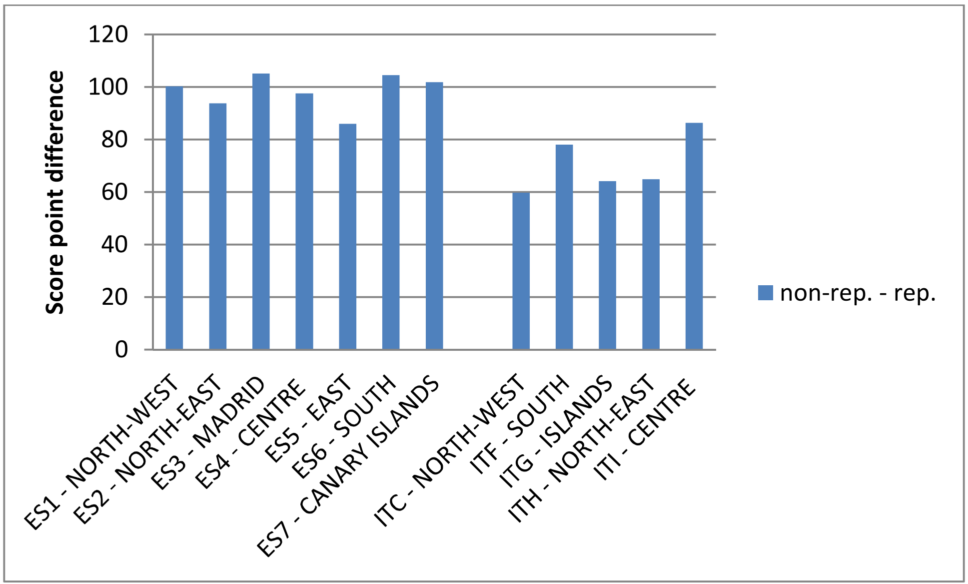
*Natives and migrants*



*Students having repeated a grade and those who have not*



KJ-NA-28868-EN-N

## JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

**EU Science Hub**
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/495702

ISBN 978-92-79-76296-3