SPANISH REPORT
Secondary Analysis

PISA 2012
Computer Based Assessment (CBA)

Problem Solving, Maths and Reading

PRELIMINARY VERSION

www.mecd.gob.es/inee

SPANISH REPORT

SECONDARY ANALYSIS

PISA 2012: Computer Based Assessment: Problem Solving, Maths and Reading

GOBIERNO DE ESPAÑA MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE
SECRETARÍA DE ESTADO DE EDUCACIÓN, FORMACIÓN PROFESIONAL Y UNIVERSIDADES
DIRECCIÓN GENERAL DE EVALUACIÓN Y COOPERACIÓN TERRITORIAL
inee Instituto Nacional de Evaluación Educativa

ÍNDICE

1. Determinants of Changes in Education Quality In Spain

Paul Cahu[1]

European Investment Bank

Luis Díez[2]

University of Minnesota

Lucas Gortázar[3]

World Bank

Abstract

We study the evolution and the determinants of education quality through time for the case of Spain. We do this by decomposing the evolution of PISA test scores, considering both average performance and quantiles of the distribution, using the unconditional quantile regression methodology proposed by Firpo, Fortin and Lemieux (2009). Doing so, we first address the impact of the economic crisis on the education system, by analyzing changes between 2009 and 2012. Although average performance slightly increased in science and reading and remained constant in math, we find that the crisis had a significant negative

[1] The author's views and opinions are personal and do not reflect the ones from the institution

[2] The author is grateful to the Bank of Spain for financial support.

[3] The author's views and opinions are personal and do not reflect the ones from the institution.

impact on students' average performance, mainly through job loss for fathers, and more significant effects for median and lowest achievers. Overall, we find that students were not equally affected by the crisis and this translated into different gains for top, median and low achievers. Second, we look at underlying explaining factors of education quality by analyzing changes in math performance between 2003 and 2012 considering attitudes toward math, motivations, school environment and teaching practices. We find a positive impact of math practices of teachers and students. Overall, results suggest a decrease in efficiency of the education system between 2003 and 2009, and an increase between 2009 and 2012, so that an improvement of the socio-economic conditions may imply an overall increase in scores on future PISA rounds, especially the lowest performers. Moreover, the study contributes to the education policy debate in Spain by bringing a modern technique for decomposing scores for different types of students, instead of the average student. This technique could be used to address relevant policy questions for the Spanish education system such as the impact of repetition in performance, the immigrants' performance gap, the differences between public and private schools, or the regional differences in scores.

INTRODUCTIÓN

In the recent years, Spain has been hit by a severe economic crisis that has undermined its capacity to provide high-quality public services to its citizens. From 2008 to 2013, Spain experienced the largest increase of unemployment rate among EU countries, which reached 24.1% in March 2012, adding pressure and difficulties on millions of households. The education sector may have not escaped from this deterioration of the economic conditions in the country, both at home and in school environment.

The fifth round of the PISA (Programme for International Student Assessment) survey was rolled out in March 2012 (following waves in 2000, 2003, 2006, and 2009). Results have shown no changes in math from the previous edition in 2009, and slight but not significant improvements in reading and science. In a broader perspective, the quality of education did not increase on average when measuring average scores since 2000. On the other hand, the PISA 2012 edition is the second assessment focused on math (repeating the 2003 edition), it allows to analyze the evolution of math performance in depth since 2003, through a richer set of background questions to students and principals.

Table 1.1. Evolution of Average Scores by scale

Test	Year				
	2000	**2003**	**2006**	**2009**	**2012**
Reading	493	481	461	481	488
Math	476	485	480	483	484
Science	491	487	488	488	496

Source: PISA 2009 and 2012 student scores for each scale.

Since the first PISA survey in 2000, the Spanish education system experienced dramatic changes: it expanded and reached universal coverage of pre-school[4]; expenditures

[4] Average years of pre-school education increased from 3.2 years in 2001-2002 to 3.8 years in 2011-12 (Ministerio de Educacion, Cultura y Deporte).

per pupil were raised significantly at least until 2010[5]; and benefited from a significant improvement of economic and social conditions till the eve of the great recession in 2009. However in spite of these positive trends the quality of the education system (as measured by PISA) did not seem to benefit from these improvements, as average scores stagnated. Understanding the underlying reasons is crucial before making further policy decisions that aim at improving the quality of the system.

In the past, some studies have analyzed the factors influencing student performance variation in Spain using PISA data from this literature, the performance of Spanish students in the most recent PISA assessment can be characterized by (i) high differences between native and immigrant students, which account for an increasing share of the school population (Zinovyeva et Al, 2013)[6]; (ii) a strong relation between student repetition (with Spain being one for the countries with higher repetition rates in OECD members) and test scores (Garcia Perez et al., 2011)[7]; (iii) differences of performance between public and private schools (Calero and Escardibul, 2007) fully explained by socio-economic differences of school population and public schools being more efficient in their use of resources; and (iv) regional differences (Garcia Perez et al., 2012) caused by differences in student's characteristics but also on how the regional education system operates, indicating room to mitigate these differences with adapted education policies.

International comparative studies have analyzed differences in PISA scores, with methods that decompose these average changes by explanatory and non-explanatory factors. The scope of study has been broad, with some studies having focused on differences across countries, types of schools or student groups[8] and others comparing the same system unit across years[9]. All these studies have analyzed differences at the mean, using Oaxaca-Blinder decompositions on linear regressions (OLS). But new methodological options have been emerging in the latest years. The labor economics literature has extensively analyzed determinants of wage gaps and wage inequality by computing counterfactual decompositions throughout the whole wage distribution - Di Nardo et al. (1996), Machado & Mata (2001) and Firpo, Fortin, Lemieux (2009) - with the underlying assumption that factor changes do not affect similarly all students. There is in fact no reason to assume that things are different in the case of education production functions. In fact, the Fortin, Firpo and Lemeiux (FFL) estimator has recently been used by Lounkaew (2013) to decompose differences between rural and urban student performance in Thailand using PISA 2009 data this method adds simplicity to Di Nardo (1996) as it allows computing unconditional distributions, which are easier to interpret in terms of marginal treatment effects of variables.

This study builds on the FFL approach, and aims at identifying and quantifying the main drivers of changes in education quality in Spain over the last decade. The methodology used will allow answering two entangled questions:

[5] In Spain, average expenditure per pupil as a percentage of GDP per capita across levels was 25 percent in 2000, 26 percent in 2003, 29 percent in 2008, and 30 percent in 2010 (OECD Education at a Glance).
[6] The authors find that most of the achievement gap between immigrants and native can be explained by individual and family characteristics)
[7] The authors find that the effect on scores of repeating is much smaller (-10% of non-repeaters average) than the counterfactual reduction that non-repeaters would suffer had they been retained as repeaters (-24% of their average).
[8] See Zinovyeva et al. (2013) and Amermueller (2004).
[9] See Barrera-Osorio et al. (2011).

1. What has been the impact of the Great Recession on students' performance, especially through changes in socio-economic factors and rising inequalities[10] between 2009 and 2012?

2. To which extent did the overall education system improve since 2003, especially through the impact of better teachers' practices and improved school environment on math performance?

The impact of an economic recession on education performance has been primarily addressed in the literature through its impact on parental labor market outcomes. Recent evidence suggests a negative relation between job loss of father and test scores and no significant effect from job loss of mother. Rege et al (2010) find that in Norway, test scores drop by 0.06 standard deviations when the father loses its job. Valenzuela (2013) tests this hypothesis for the case of job loss in Catalonia (Spain) during the great recession in the period 2008-2012 and finds a causal negative relationship between father's job loss and student performance, with a marginal effect between 0.13 and 0.16 standard deviations of the test score (and negative but not significant effect for mother's job loss), but with larger impact for disadvantaged students. This author also finds that father's motivation drops after the job loss, being the key driver in the fall in student scores, suggesting potential parental conflicts at home and changes in role perceptions on the child, especially boys.

Beyond socioeconomic conditions, which for the major part cannot be influenced by education policy in the medium run, math performance is also affected by several factors, which can be either reinforced or smoothened by teachers' practices. The 2012 and 2003 PISA surveys control for some of these factors, including: (i) math anxiety, (ii) instrumental motivation, (iii) math self-efficacy; (iv) math self-concept; (v) disciplinary climate, (vi) attitudes toward school; (vii) interest for math; (viii) feeling of belonging at school; (ix) teacher support and (x) teacher-student relationships. Math anxiety was first documented in the early 1970s (Richardson and Suinn), as an apprehension for math which impedes math performance, regardless of actual cognitive skills. Math anxiety has furthermore been linked to attitudes and expectations of teachers and parents (Zavaslvsky, 2004). A meta-analysis by Hembre (1990) confirmed that math anxiety is related to poor performance, low achievement in test and avoidance of math. Beliefs have also an important influence on math performance. Self-efficacy, which relates to the confidence of one's own ability to complete tasks, has been shown to have a positive impact on math performance for both students (Gail and Betz, 1989) and adults (Watts, 2011). And self-efficacy too can be influenced by specific classroom practices, leading to improvement in performance (Fast et al, 2014). The self-concept is a psychological notion that embeds how people self-evaluate their abilities and skills. Self-concept is affected by socialization and is related to academic performance as well. Flook et al (2005) have demonstrated for instance that peer acceptance within the class has an impact on self-concept which in turns affects academic performance several years later. The empirical entanglements between self-concept, students' socialization and academic performance justify why measures of self-concept and the feeling of belonging at school are used in the PISA surveys. Moreover, math assessments designed specifically to measure the effects of changes in teacher practices have shown that teachers working at improving their explanations tend to develop the reasoning skills of their students (Silver and Stein, 1996). Finally, disciplinary climate is also likely to affect students' performance. Although the importance of discipline

[10] According to Eurostat, the Gini Index of equivalised disposable income for Spain increased from 31.9 in end 2008 to 35 in 2012.

varies among countries and cultures, students taught in more orderly classrooms tend to learn better, see Arum & Velez (2012), Figlio (2007) or Marks (2010).

Data Description And Methodology

The Programme for International Student Assessment (PISA) is a survey coordinated by the Organization for Economic Cooperation and Development (OECD) which evaluates student competencies in reading, mathematics, and science every three years since 2000. A representative sample of 15-year old students enrolled in schools sit standardized assessments which allow estimating proficiency in reading, math and science from item response distribution. Students also fill up a comprehensive questionnaire including detailed data about their individual and family characteristics. Moreover, school principals fill up another complementary questionnaire to capture features of the school, including organization, school resources, sharing of responsibilities, policies and strategies implemented and school climate.

To address critics about reliability, OECD now recommends restricting the time comparison of a given scale to rounds of the PISA survey where full assessment of this scale has taken place[11]. A full assessment in math was conducted for the second time in 2012, following the first edition in 2003, so that both rounds should be used to look at math performance across time. However, we also consider in a first step the evolutions of scores in the three scales during the period 2009-2012, to understand how the economic recession may have affected student outcomes through observable factors. Such analysis is needed to assess the consequences of the recent economic turmoil on Spain's education system.

Re-computing OECD indexes

The PISA data sets exhibit built-in indexes intending to measure socioeconomic factors, such as indexes of home educational resources (HEDRES), family wealth (WEALTH), cultural possessions (CULTPOS), a mixture of the previous three (HOMEPOS) or a synthetized index of social, cultural, and economic status, which combines HOMEPOS with highest parental education (PARED), and an index of highest occupational status of parents (HISEI). However, these indexes are made to reflect the diversity of factors that matter among all countries of the sample and are not perfectly suitable to analyse trends within a given country. To compare precisely the evolution of students' performance between two different years, we re-computed these indexes from the basic items of the questionnaire. We use items that are common in both questionnaires, as some indexes have become more comprehensive with the years. Using similar weights to compute indexes from the basic items allows displaying more clearly the overall increase in living standards that has occurred during the last decade in Spain, both when comparing changes between 2003 and 2012, or between 2009 and 2012[12].

[11] Full assessment of a PISA tests in a specific scale increase the number of assessment items and questionnaire items, meaning that when the full assessment coincides in the same scale for two different rounds, the degree comparability along time is more reliable.

[12] Indexes were computed by calculating partial correlations of each component in a linear regression of student reading scores on the components taken together.

Labor market Variables, Family Background, and pre-primary education

Given the large changes that affected the Spanish labour market in Spain between 2009 and 2012, we look in depth at the variable of interest, the parental labor status. Although parental employment is defined more broadly in the PISA survey, we only consider here whether or not parents are fully employed. Other variables like family structure, the immigration background and the participation to pre-primary school are strong indicators of socio-economic development in the case of Spain. We also consider missing values as a separate category for the structure of family, as the structure of the household is a straightforward information that any children should be able to provide. Missing data are therefore a strong sign of family difficulties, such as the ones that occur during a divorce or separation period.

Empirical Strategy

Our approach builds on the education production functions literature. To analyze the determinants that influence of the evolution of students' performance across time, variables can be divided into three different groups: (i) individual and family factors (such as age, gender, family socio-economic status, but also student intrinsic motivation and engagement in school); (ii) school factors: through peer characteristics (coming from socio-economic variables averaged at the school level, proportion of repeaters, dropouts, or girls), geographical variables (urban household, or region), school resources, teacher characteristics, pedagogical strategies, principal characteristics, parental engagement, disciplinary climate, or quality of resources; and (iii) systemic factors: organization and ownership of schools, school autonomy and accountability, and early streaming policies. In this study, we use the following specification for the education production function:

$$Y_i = \alpha + X_i'\beta + P_i'\mu + Z_i'\delta + \epsilon_i \qquad (1)$$

Where Y_i represents the student score, X_i is a vector of student and family background characteristics, P_i is a vector of student and family characteristics averaged at the school level (to capture peer effects), Z_i is a vector of school-specific characteristics, and ϵ_i is an error term. If we compare two groups of students at different years (t_1 and t_0), the difference of PISA test performance is:

$$Y_{t1} - Y_{t0} = (\alpha_{t1} - \alpha_{t0}) + (X_{t1}'\beta_{t1} - X_{t0}'\beta_{t0}) + (P_{t1}'\mu_{t1} - P_{t0}'\mu_{t0}) + (Z_{t1}'\delta_{t1} - Z_{t0}'\delta_{t0}) + (\epsilon_{t1} - \epsilon_{t0}) \qquad (2)$$

Adding and subtracting $X_{t1}'\beta_{t1}$, $P_{t1}'\mu_{t1}$ and $Z_{t1}'\delta_{t1}$ on the right side of equation (2) and rearranging gives the following expression:

$$Y_{t1} - Y_{t0} = \{(X_{t1} - X_{t0})'\beta_{t1} + (P_{t1} - P_{t0})'\mu_{t1} + (Z_{t1} - Z_{t0})'\delta_{t1}\} +$$

$$\{X_{t1}'(\beta_{t1} - \beta_{t0}) + P_{t1}'(\mu_{t1} - \mu_{t0}) + Z_{t1}'(\delta_{t1} - \delta_{t0}) + (\alpha_{t1} - \alpha_{t0}) + (\epsilon_{t1} - \epsilon_{t0})\} \qquad (3)$$

Equation (3) is the so called Oaxaca-Blinder decomposition, where the first term in brackets of the right-hand side of the equation is the **explained** part of the differences of scores and it represents how much variation in observed characteristics of students and schools explain student scores. The second term in brackets of the right-hand side of the equation, considered the **unexplained** part, represents at the same time the differences **in structural changes** in returns of characteristics (how differences in characteristics across years are transformed into test scores) **and** all **unobserved characteristics** which the model cannot capture. Independently of which of the two empirical questions we want to answer and the variables of interest that are related, we add control variables that may be correlated with unobserved student and school characteristics in order to minimize the impact of omitted variables which may induce biases in the statistic of interest.

Following Lounkaew (2013), this study goes beyond the Oaxaca-Blinder approach at the mean (with linear regression models) and extends to other statistics of the distribution. In our case, the rural-urban gap is replaced by a year t_1 with year t_0 gap. Usually, unconditional quantile regressions are not feasible given that the law of iterated expectations does not apply for quantiles. In order to make this unconditional counterfactual exercise possible, Firpo, Fortin, and Lemieux (2009) propose a technique based on re-centered influence functions (RIF) of a quantile of interest q_τ:

$$RIF(I; q_\tau) = q_\tau + \frac{\tau - D(I \leq q_\tau))}{f_I(q_\tau)} \quad (4)$$

where D is an indicator function, $f_I(.)$is the density of the marginal distribution of scores. In practice, $RIF(I; q_\tau)$ is not observed, so that its sample counterpart is:

$$RIF(I; \hat{q}_\tau) = \hat{q}_\tau + \frac{\tau - D(I \leq \hat{q}_\tau))}{\hat{f}_I(q_\tau)} \quad (5)$$

where $\hat{q}_\tau$ is the sample quantile and $\hat{f}_I(q_\tau)$ is the kernel density estimator. A crucial characteristic of this technique is that it provides a simple way of interpreting the marginal impact of an additional unit of certain factor in student PISA scores. Once the unconditional quantile regression has been computed for different quantiles of the distribution, we decompose the results following the Oaxaca-Blinder approach.

The Impact of the Great Recession on Student Performance

This section analyzes changes in observable student and school characteristics (including crisis-related and other observable factors) in the PISA sample for Spain and how this impacted scores in math, reading and science. Our production function follows (3) and divides student and school characteristics by "Crisis" and "Non-Crisis" group of variables, so that the resulting production function builds on (1):

$$Y_i = \alpha + X'_{NCi}\beta_{NC} + P'_{NCi}\mu_{NC} + Z'_{NCi}\delta_{NC} + X'_{Ci}\beta_C + P'_{Ci}\mu_C + Z'_{Ci}\delta_C + \epsilon_i$$
(5)

and distinguishes student variables at the household, student variables at the school level, and school specific variables between two categories: non-crisis variables (NC) and crisis variables (C).

These characteristics are displayed in Table 1.2, where we can see average evolutions of student and school covariates from 2009 to 2012. As said previously, we create comparable indexes from the PISA database. Variables that changed significantly during the crisis years are the share of single parents and the share of students unable to report their family structure. Most importantly, variables directly affected by the economic crisis do in fact change. Father's full time employment fell from 73% in 2009 to 67% in 2012. Moreover, home educational resources and cultural possessions became scarcer, while the wealth index improved, traducing better endowments in durable goods. School variables related to shortage of teachers, equipment, and materials, did not experience major changes.

Table 1.2. Student and School Characteristics, 2009-2012

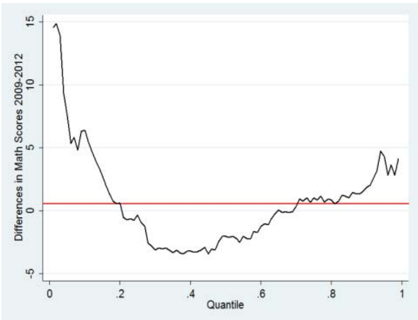
Variable	2009 Obs.	2009 Mean	2009 Std. Dev.	2012 Obs.	2012 Mean	2012 Std. Dev.
Individual Non-Crisis						
Age	25887	15.85	0.29	25313	15.86	0.29
Grade	25887	9.53	0.68	25313	9.56	0.67
Girl	25887	0.49	0.50	25313	0.49	0.50
Foreign Language	25513	0.18	0.38	24686	0.19	0.39
Parental Highest Education (Year)	25257	12.06	3.87	24515	11.36	3.96
Immigrant First generation	25887	0.08	0.28	24268	0.08	0.28
Immigrant Second Generation	25887	0.01	0.10	24268	0.01	0.12
Pre-school 1 year	25351	0.09	0.28	24934	0.08	0.28
Pre-school +2 years	25351	0.87	0.34	24934	0.86	0.35
Single Parent Family	25887	0.13	0.34	25313	0.10	0.29
Missing Family Information	25887	0.01	0.10	25313	0.07	0.25
Individual Non-Crisis (School Average)						
Immigrant First generation (School)	25887	0.01	0.02	25313	0.01	0.03
Immigrant Second Generation (School)	25887	0.08	0.10	25313	0.08	0.11
Pre-School +2 years (School)	25887	0.87	0.10	25313	0.86	0.10
Foreign Language (School)	25887	0.18	0.26	25313	0.19	0.27
Percentage of Girls (School)	24754	0.50	0.09	25313	0.46	0.14
Highest Parental Education (School)	25887	12.05	1.89	25313	10.95	1.92
Grade (School)	25887	9.53	0.26	25313	9.56	0.26
Individual Crisis						
Father full-time employed	25887	0.73	0.44	25313	0.68	0.47
Mother full-time employed	25887	0.44	0.50	25313	0.45	0.50
Index of Highest Parental Occupation	24843	45.85	17.13	24906	46.95	21.51
Index of Home Educational Resources	25481	-0.13	0.91	24379	-0.19	0.90
Index of Home Cultural Posessions	25512	0.19	0.87	24865	0.11	0.88
Index of Wealth	25456	-0.21	0.76	25082	-0.14	0.71
Individual Crisis (School Average)						
Father full-time employed (School)	25887	0.73	0.11	25313	0.68	0.13
Mother full-time employed (School)	25887	0.44	0.14	25313	0.45	0.16
Index of Highest Parental Occupation (School)	25813	45.78	9.08	25313	46.24	10.52
Index of Home Educational Resources (School)	25887	-0.13	0.26	25313	-0.19	0.25
Index of Home Cultural Resources (School)	25887	0.19	0.32	25313	0.11	0.33
Index of Wealth (School)	25887	-0.21	0.33	25313	-0.14	0.27
School Resouces-Crisis						
Quality of School Resources	25251	0.01	0.85	25009	0.02	0.86
Math Teacher Shortage	25212	1.06	0.30	24955	1.13	0.40
Science Teacher Shortage	25212	1.08	0.32	24933	1.11	0.37
Reading Teacher Shortage	25144	1.07	0.30	24979	1.09	0.34
Shortage of Lab Equipment	25165	2.09	0.88	24808	2.01	0.91
Shortage of Library Materials	25193	2.13	0.86	24907	1.93	0.84
Shortage of Instruction Materials	25148	1.44	0.67	24956	1.44	0.67
School Factors						
Rural	25887	0.29	0.46	25087	0.28	0.45
Public	25887	0.64	0.48	25313	0.67	0.47
Disciplinary Climate at School	25887	0.08	0.36	25309	-0.04	0.43

Note: Data from PISA 2009 and 2012.

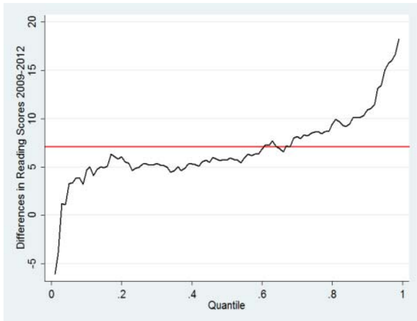
Respective impacts of the variables displayed in Table 1.2 on student scores is documented in Table 1.A1 in the Annex, using linear regressions for all three scales (math, reading, and science) in separate years 2009 and 2012, following model (5). Results display the importance of grade repetition, gender, immigration status, pre-school education, family structure (including students unaware of their family structure), both at the individual and averaged at the school level. Regarding variables related to wealth and income, it can be seen that father's employment has more explanatory power than mother's, and that home educational resources have a very significant impact on performance, for both years and all scales.

Figura 11 illustrates differences in scores between 2009 and 2012 for each scale across the distribution of achievers. In math, a stable average score actually hides heterogeneous evolutions at different quantiles of the distribution. While students from bottom quantiles increased their performance from 5 to up to 15 points, median achievers experienced a slight drop in scores, inducing declining variation in math performance since 2009. In reading, performance of median achievers increased by 7 points, but lowest achievers improved less than average or even declined, while top performers improved beyond average gains, by up to 15 points. Finally changes in science were more homogeneous with an overall improvement by 8 points for all but the low performers, who gained even more, up to 12 points. This marked heterogeneity in how of performance evolved, especially at the queues of the distribution justifies the use of unconditional quantile regressions in the analysis of determining factors.

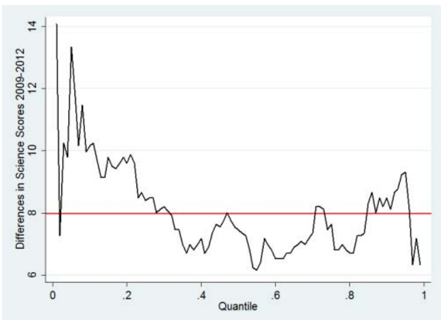
Figure 1. Change in Score by quantiles in each scale between 2009 and 2012



Differences in Math Scores



Differences in Reading Scores



Differences in Science Scores

Fuente: PISA 2009 y 2012 resultados para cada disciplina la línea roja representa el cambio promedio.

Table 1.3 shows the Oaxaca-Blinder decomposition for math, by grouping different variable following equation (5). A detailed description of the effect of each variable can be seen in the Annex. We find an average negative impact of 6.7 points by explanatory variables, and a positive impact of unexplained variables, suggesting an increase of efficiency of the education system in a period of deterioration of living standards.

Moreover, the impact of crisis-related individual factors (both individual and school average) is not negligible and accounts for most of the explained part changes. This can be interpret as a counterfactual treatment of the crisis-related factors into PISA performance, so that the absence of this factors would have implied an average positive net effect which would have pushed the math scores up around 6 points in the scale. The impact is less smaller for percentiles 10 and 90, but stronger at the center of the distribution. Detailed analysis (see

7	Math		Read		Science	
	2009	2012	2009		2009	2012
Individual Non-Crisis				*Individual Non-Crisis*		
Age	-2.161	0.187	5.397**	Age	-2.161	0.187
	(2.503)	(2.289)	(2.412)		(2.503)	(2.289)
Grade	63.97***	60.75***	57.56***	Grade	63.97***	60.75***
	(1.319)	(1.587)	(1.237)		(1.319)	(1.587)
Girl	-32.09***	-25.77***	15.96***	Girl	-32.09***	-25.77***
	(1.448)	(1.357)	(1.350)		(1.448)	(1.357)
Foreign Language	-1.167	-2.612	-5.138**	Foreign Language	-1.167	-2.612
	(3.645)	(2.448)	(2.489)		(3.645)	(2.448)
Parental Highest Education (Years)	0.197	0.514**	0.182	Parental Highest Education (Years)	0.197	0.514**
	(0.275)	(0.236)	(0.240)		(0.275)	(0.236)
Immigrant First Generation	-18.29***	-11.72***	-19.85***	Immigrant First Generation	-18.29***	-11.72***
	(3.212)	(2.800)	(2.662)		(3.212)	(2.800)
Immigrant Second Generation	-14.27**	-9.916**	-8.413	Immigrant Second Generation	-14.27**	-9.916**
	(6.983)	(4.977)	(6.163)		(6.983)	(4.977)
Pre-School 1 year	2.499	8.306**	4.691	Pre-School 1 year	2.499	8.306**
	(4.402)	(3.658)	(4.880)		(4.402)	(3.658)
Pre-School 2+ Years	16.13***	20.38***	13.82***	Pre-School 2+ Years	16.13***	20.38***
	(3.924)	(2.884)	(4.077)		(3.924)	(2.884)
Single Parent Family	-1.034	6.010**	0.785	Single Parent Family	-1.034	6.010**
	(2.281)	(2.427)	(2.251)		(2.281)	(2.427)
Family Missing Information	-36.21***	-13.85***	-32.38***	Family Missing Information	-36.21***	-13.85***
	(7.919)	(3.149)	(8.054)		(7.919)	(3.149)
Individual Non-Crisis (School Average)				*Individual Non-Crisis (School Average)*		
Immigrant Second Generation (School)	72.00	-21.21	109.2*	Immigrant Second Generation (School)	72.00	-21.21
	(62.54)	(63.00)	(55.76)		(62.54)	(63.00)
Immigrant First Generation (School)	6.396	30.17**	23.25	Immigrant First Generation (School)	6.396	30.17**
	(14.70)	(14.18)	(14.98)		(14.70)	(14.18)
Pre-School 2+ Years (School)	-6.312	11.77	28.30**	Pre-School 2+ Years (School)	-6.312	11.77
	(14.80)	(12.79)	(13.44)		(14.80)	(12.79)
Foreign Language (School)	4.544	-8.411	0.548	Foreign Language (School)	4.544	-8.411
	(7.547)	(7.331)	(7.980)		(7.547)	(7.331)
Percentage of Girls in School	60.40***	-8.704	49.66	Percentage of Girls in School	60.40***	-8.704
	(20.41)	(11.26)	(33.07)		(20.41)	(11.26)
Parental Highest Education Years (School)	4.129**	1.959	3.077*	Parental Highest Education Years (School)	4.129**	1.959
	(1.713)	(1.535)	(1.585)		(1.713)	(1.535)
Average Grade in School	-5.780	7.593	4.922	Average Grade in School	-5.780	7.593
	(7.518)	(6.465)	(7.676)		(7.518)	(6.465)

Individual Crisis			
Father Full Time Employed	2.832*	2.628*	0.866
	(1.534)	(1.499)	(1.541)
Mother Full Time Employed	1.998	0.392	-0.364
	(1.290)	(1.372)	(1.270)
Index of Highest Parental Occupation	0.411***	0.378***	0.408***
	(0.0504)	(0.0448)	(0.0527)
Index of Home Educational Resources	3.035***	1.906**	2.195**
	(0.915)	(0.883)	(0.874)
Index of Cultural Possessions	9.943***	10.67***	10.87***
	(0.875)	(0.960)	(0.793)
Index of Wealth	-2.292**	-2.029*	-4.276***
	(1.059)	(1.222)	(1.056)
Individual Crisis (School Average)			
Father Full Time Employed (School)	55.09***	51.77***	7.333
	(14.45)	(12.97)	(14.40)
Mother Full Time Employed (School)	14.63	9.069	34.64***
	(12.44)	(10.81)	(11.80)
Index of Highest Parental Occupation (School)	-0.285	-0.112	-0.00753
	(0.382)	(0.276)	(0.377)
Index of Home Educational Resources (School)	*1.666*	*1.971*	*5.524*
	(6.493)	(5.551)	(6.401)
Index of Cultural Possessions (School)	28.70***	15.32**	18.77**
	(7.590)	(6.184)	(7.432)
Index of Wealth (School)	-13.93***	-10.82*	-9.152*
	(5.259)	(5.665)	(5.143)
School Resources-Crisis			
Quality of School Resources	-1.906	-2.259	-0.690
	(2.973)	(2.039)	(2.548)
Shortage of subject teacher	2.444	-2.702	0.927
	(3.038)	*(2.993)*	*(3.416)*
Shortage of Library Materials	-4.449*	-2.916	-1.799
	(2.390)	(1.964)	(2.301)
Shortage of Instruction Materials	-1.571	-2.012	0.374
	(2.362)	(2.362)	(2.108)
School Factors			
Rural	7.239**	1.693	-0.140
	(3.308)	(2.949)	(3.399)
Public School	5.939*	-0.520	1.533
	(3.414)	(3.088)	(3.366)
Index of Disciplinary Climate at School	10.20**	3.411	8.748*

Individual Crisis		
Father Full Time Employed	2.832*	2.628*
	(1.534)	(1.499)
Mother Full Time Employed	1.998	0.392
	(1.290)	(1.372)
Index of Highest Parental Occupation	0.411***	0.378***
	(0.0504)	(0.0448)
Index of Home Educational Resources	3.035***	1.906**
	(0.915)	(0.883)
Index of Cultural Possessions	9.943***	10.67***
	(0.875)	(0.960)
Index of Wealth	-2.292**	-2.029*
	(1.059)	(1.222)
Individual Crisis (School Average)		
Father Full Time Employed (School)	55.09***	51.77***
	(14.45)	(12.97)
Mother Full Time Employed (School)	14.63	9.069
	(12.44)	(10.81)
Index of Highest Parental Occupation (School)	-0.285	-0.112
	(0.382)	(0.276)
Index of Home Educational Resources (School)	*1.666*	*1.971*
	(6.493)	(5.551)
Index of Cultural Possessions (School)	28.70***	15.32**
	(7.590)	(6.184)
Index of Wealth (School)	-13.93***	-10.82*
	(5.259)	(5.665)
School Resources-Crisis		
Quality of School Resources	-1.906	-2.259
	(2.973)	(2.039)
Shortage of subject teacher	2.444	-2.702
	(3.038)	*(2.993)*
Shortage of Library Materials	-4.449*	-2.916
	(2.390)	(1.964)
Shortage of Instruction Materials	-1.571	-2.012
	(2.362)	(2.362)
School Factors		
Rural	7.239**	1.693
	(3.308)	(2.949)
Public School	5.939*	-0.520
	(3.414)	(3.088)
Index of Disciplinary Climate at	10.20**	3.411

				School		
	(4.545)	(3.062)	(4.476)		(4.545)	(3.062)
Constant	175.9***	158.2***	21.88	Constant	175.9***	158.2***
	(51.16)	(46.31)	(54.59)		(51.16)	(46.31)
Observations	22,316	21,749	22,283	Observations	22,316	21,749
R-squared	0.461	0.451	0.454	R-squared	0.461	0.451

in Annex) shows that most of the effect is linked to father losing their job and fewer cultural possessions (most important at percentiles 30, 50, and 70). Moreover, there was no significant impact of school resources which may have dampen performance as a consequence of reduced public funding. Also, shifts in family structure had significant marginal negative impact.

In fact, different expositions to the crisis resulted in heterogeneous changes in performance. Low achievers who are on average living in more difficult socio-economic conditions were twice as less affected by crisis-related factors than median or top/median achievers (at the 70^{th} percentile). This finding is consistent with the fact that fathers of disadvantaged children were likely to be unemployed or part-time employed already before the crisis. Students from the center of the distribution, on the contrary, took the largest hit, as their fathers were directly exposed on the labor market. Top achievers were affected to a lesser extent due to little job losses by parents. These differentiated impacts drove the overall heterogeneity in math performance change, with lowest achievers achieving sizeable gains in math performance while median and top achievers did not, as the unexplained part of the decomposition is overall homogenous for all students. It is to be noted that this profile remains consistent for the three scales (see Tables 1.3, 1.4 and 1.5).

Table 1.3. Math Scale: Decomposition of Change in Scores in 2009-2012 by mean and quantiles

Dependent Variable: Math	Mean	P 10	P 30	P 50	P 70	P 90
Explained (1)	-6.768**	-2.542	-8.201	-7.927*	-10.46**	-7.454***
	(3.273)	(2.930)	(5.021)	(4.145)	(5.273)	(2.721)
Individual Non-Crisis	0.850	0.470	1.743	1.189	1.219	0.586
	(1.362)	(1.359)	(2.345)	(1.677)	(1.761)	(0.747)
Individual Non-Crisis (Peer Effects)	-1.914	0.288	-1.849	-2.201	-3.415	-2.627
	(1.779)	(1.868)	(3.000)	(2.348)	(3.105)	(1.671)
Individual Crisis	-0.739	-0.706*	-1.228*	-0.802	-1.142	-0.611
	(0.572)	(0.407)	(0.684)	(0.694)	(1.098)	(0.597)
Individual Crisis (Peer Effects)	-4.923***	-2.913**	-6.926***	-5.877***	-6.989***	-4.362***
	(1.230)	(1.241)	(1.950)	(1.647)	(2.183)	(1.271)
School Resources Crisis	-0.460	-0.550	-0.578	-0.562	-0.739	-0.726
	(0.400)	(0.425)	(0.648)	(0.560)	(0.737)	(0.461)
School Non-Crisis	0.418	0.869	0.637	0.325	0.601	0.287
	(0.482)	(0.570)	(0.766)	(0.611)	(0.908)	(0.466)
Unexplained (2)	10.03***	10.58***	11.04***	10.47***	16.56***	11.36***
	(2.813)	(2.804)	(4.202)	(3.419)	(4.757)	(2.787)
Difference (1+2)	3.262	8.035***	2.839	2.539	6.095	3.901
	(3.221)	(2.818)	(4.006)	(3.355)	(4.420)	(2.482)
Observations	44,065	44,065	44,065	44,065	44,065	44,065

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1. Variables effects are grouped after a detailed decomposition as shown in Annex.

.

Table 1.4 displays similar analysis results for reading. The average impact of crisis related factors is smaller than for math, but remains significant. Marginal impacts are moreover larger for median achievers. Compared to math, the marginal impact of parental job loss on reading is however smaller, although cultural possessions tend to matter as much as for math. The unexplained part of the Oaxaca-Blinder decomposition contributes more significantly to the change in performance than for math, especially for top performing students, traducing the fact that the main drivers of reading performance causes remain unknown. The unexplained factors' net effect is even very positive and significant and again suggests an overall improvement of efficiency of the system, especially for top-performing students.

Table 1.5 displays similar results for science (see detailed results in Annex). As in the case of math and reading, median achievers were more affected by the crisis. As for reading and math, changes in cultural possessions and father full-time employment appear as the main observable drivers of the evolution in performance. But as in the case of reading the unexplained part of the evolution was biased toward median and top achievers, explaining why overall, low achievers gained less.

The decomposition of changes in students' performance during the crisis years reveal that the economic crisis did not affect all students similarly and that the school system benefitted from an overall improvement which counteracted the effects of the crisis but was not homogenous either for students of different proficiency levels. To investigate the causes of this shift in quality, we now exploit the comparability of the 2003 and 2012 surveys to quantify the impact of potential gains in math in terms of teachers' practices and school learning environment.

Table 1.4 Reading Scale: Decomposition of Change in Scores in 2009-2012 by mean and quantiles

Dependent Variable: Reading	Mean	P 10	P 30	P 50	P 70	P 90
Explained (1)	-2.972	0.395	-4.495	-2.250	-3.295	-4.856**
	(3.404)	(2.749)	(4.553)	(3.354)	(2.990)	(2.013)
Individual Non-Crisis	0.871	1.039	1.749	0.880	0.757	-0.352
	(1.264)	(1.099)	(1.813)	(1.248)	(0.961)	(0.492)
Individual Non-Crisis (Peer Effects)	-0.680	1.583	-2.013	0.00742	-1.463	-2.196
	(2.182)	(1.939)	(3.336)	(2.212)	(1.903)	(1.451)
Individual Crisis	-1.626***	-1.099***	-2.130***	-1.419***	-1.462**	-0.549
	(0.554)	(0.361)	(0.648)	(0.501)	(0.625)	(0.447)
Individual Crisis (Peer Effects)	-1.614	-1.576	-2.304	-1.588	-1.447	-1.497
	(1.575)	(1.217)	(2.213)	(1.723)	(1.543)	(0.988)
School Resources Crisis	0.222	0.272	0.445	-0.106	0.574	-0.120
	(0.526)	(0.489)	(0.746)	(0.526)	(0.497)	(0.350)
School Non-Crisis	-0.145	0.176	-0.242	-0.0252	-0.255	-0.141
	(0.502)	(0.445)	(0.683)	(0.490)	(0.533)	(0.328)
Unexplained (2)	12.09***	5.819*	9.951**	7.647**	12.96***	16.12***
	(3.214)	(3.011)	(4.809)	(3.762)	(3.097)	(2.697)
Difference (1+2)	9.117***	6.213**	5.456	5.396	9.670***	11.26***
	(3.316)	(2.695)	(4.581)	(4.242)	(3.305)	(2.643)
Observations	44,052	44,052	44,052	44,052	44,052	44,052

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1. Variables effects are grouped after a detailed decomposition as shown in Annex.

Table 1.5 Science Scale: Decomposition of Change in Scores in 2009-2012 by mean and quantiles

Variables	Mean	P 10	P 30	P 50	P 70	P 90
Explained (1)	-5.329	-1.883	-3.240	-6.653	-8.659*	-10.68**
	(3.379)	(2.044)	(3.701)	(4.605)	(5.160)	(4.145)
Individual Non-Crisis	0.202	0.177	0.557	0.338	0.475	-0.431
	(1.193)	(0.843)	(1.383)	(1.482)	(1.540)	(1.096)
Individual Non-Crisis (Peer Effects)	-1.085	0.545	-0.353	-1.278	-2.327	-3.896
	(2.190)	(1.449)	(2.462)	(3.183)	(3.209)	(2.546)
Individual Crisis	-1.215**	-0.421	-1.430***	-1.487**	-1.932**	-1.448
	(0.545)	(0.261)	(0.474)	(0.664)	(0.958)	(0.932)
Individual Crisis (Peer Effects)	-3.504**	-2.550***	-2.199	-4.458**	-5.180**	-4.984**
	(1.398)	(0.894)	(1.533)	(1.988)	(2.207)	(1.954)
School Resources Crisis	-0.254	-0.203	-0.0151	-0.158	-0.700	-1.118
	(0.536)	(0.337)	(0.664)	(0.678)	(0.875)	(0.752)
School Non-Crisis	0.528	0.569	0.199	0.390	1.004	1.200
	(0.476)	(0.420)	(0.508)	(0.637)	(0.768)	(0.784)
Unexplained (2)	16.10***	10.89***	13.67***	17.20***	21.23***	22.26***
	(3.417)	(2.890)	(3.851)	(4.451)	(4.861)	(4.060)
Difference (1+2)	10.77***	9.008***	10.43***	10.55***	12.57***	11.59***
	(3.189)	(2.864)	(3.519)	(3.672)	(3.718)	(3.431)
Explained (1)	43,924	43,924	43,924	43,924	43,924	43,924

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1. Variables effects are grouped after a detailed decomposition as shown in Annex.

DRIVERS OF MATH PERFORMANCE BETWEEN 2003 AND 2012

In this section we measure the impact of math teaching and learning practices on math scores and if changes in these reported skills between 2003 and 2012 can explain changes in scores during this same period of time. The PISA surveys measure cognitive skills in reading, math and science. Although achievements on those three scales are highly correlated, skills surveys in different countries have pointed out that math skills are more critical for employment than literacy or science, see for instance BIS (2012) for England, World Bank (2013) for Russia or results from the OECD PIAAC survey (2013).

We build a series of indicators which measure student behaviors and beliefs, and related teacher practices with math subject to estimate the importance of overall teacher's effectiveness on math scores. The model we use follows model (1) in previous section, but adds a separate group of variables related to teaching practices and the overall learning environment so that:

$$Y_i = \alpha + X_i'\beta + P_i'\mu + Z_i'\delta + M_i'\eta + \epsilon_i \qquad (6)$$

where M_i represents a group of indexes and variables measuring student behaviors and beliefs, teaching practices and school environment relevant for the study of math performance.

Creating indicators of math practices and school climate

Using most of the information available in the PISA data set, we build indexes related to the quality of the school environment, the behaviours of students and the practices of teachers as a set of measures of the overall teaching and learning practices within the schools. It is obvious that values of these indexes are likely to be affected by intrinsic characteristics of students, teachers and school directors. Therefore, they are not purely measuring practices of teachers and principals – which may be improved through targeted interventions – but reflect also factors that are given and may not be changed in the medium run. Moreover, self-reported answers related to character skills are likely to suffer from reference biases, given that these questions mean different things to different people. This section argues, however, that computing a comprehensive set of indicators introduced along these measures of overall school practices (and controlling for social, economic and cultural factors at both the individual and the school level at the same time) is likely to reduce substantially any biases of these measures.

Both PISA data sets from 2003 and 2012 display built-in indicators on attitudes, motivations, and practices, mostly but not all, related to math learning. But to ensure a strict comparability of the two data sets, we first only use the indicators that appear on both data sets. Second, we re-compute the indicators using the raw answer to the basic questionnaire for the questions asked that are identical in both rounds. This process allows us to obtain measurement of ten different dimensions of what happens within the school (see Table 1.6). In fact, the items used to measure the "support from teachers" and the quality of the "teacher-student" relationship are quite similar, and the correlation between the two indicators is about 0.45. These two indicators therefore measure similar behaviours and the items related in the questionnaire will be treated as part of the same dimension of practices.

Table 1.6 Dimensions of school practices in PISA 2003 & 2012

Dimensions	Definition	Items in 2003 questionnaire	Items in 2012 questionnaire
Anxiety related to Math	Stress when practicing of math	Q32: 1,3,5,8,10	Q42: 1,3,5,8,10
Disciplinary climate	Quality of discipline	Q38: 2,6,8,9,11	Q81: 1-5
Math Self-Efficacy	Self-confidence toward doing specific math tasks	Q31: 1-8	Q37: 1-8
Math self-concept	Self-evaluation of overall math proficiency	Q32: 2,4,6,7,9	Q42: 2,4,6,7,9
Support from teachers	Attitudes of teachers toward students	Q38: 1,3,5,7,10	Q77: 1,2,4,5,6
Teacher-student relationship		Q26: 1-5	Q86: 1-5
Attitude toward school	Beliefs about usefulness of school	Q24: 1,-4	Q88: 1-4
Intrinsic motivation	Beliefs about usefulness of math	Q30: 2,5,7,8	Q29: 2,5,7,8
Interest for math	Like doing math	Q30: 1,3,4,6	Q29: 1,3,4,6
Belong to school	Feeling that belong to school	Q27: 1-6	Q87: 1-6

Source: PISA 2003 and 2012.

Basic indicators are built as the first principal component over a series of responses of items over a Likert scale, as usual with psychometric measurement. The principal component is computed as the arithmetic average of the responses.[13] Interestingly, the correlation between the indicator and students' performance in math is not always the one which can be suspected (see Table 1.8 below). We would expect for instance a positive impact of math attitudes, motivations and practices over student scores. Nevertheless, four out of ten indicators do not correlate with performance in the suspected direction. It is to be noted that adding controls for individual factors does not change this fact.

[13] This equal weighting is indeed the one which maximizes the likelihood of the observations.

Table 1.7 Marginal impact of school practices Indexes on average math score, 2003-2012 pooled PISA data

Variable	Mean estimate	T-stat	Expected sign
Anxiety related to Math	-12.7	14.1	YES
Disciplinary climate	9.4	12.0	YES
Math Self-Efficacy	34.5	42.5	YES
Math self-concept	15.9	16.4	YES
Support from teachers	-7.6	10.6	NO
Teacher-student relationship	-4.2	5.8	NO
Attitude toward school	4.8	7.0	YES
Intrinsic motivation	7.1	9.9	YES
Interest for math	-9.5	10.1	NO
Belong to school	-3.9	5.6	NO

Source: Author's calculations. No other controls. Standard deviations are computed by clustering observations by schools. Number of observations: 17,867. Adjusted R^2 is 0.31. OLS regressions.

The negative correlation between the attitudes of teachers and students' performance may be due to an endogeneity bias: teachers may have to do more efforts when confronted with less proficient students. To test that possibility, we now look at the marginal impact of the underlying items used to build the "support from teachers" and the "teacher-student relationship" indicators, the two most policy-related indicators of the table. Table 1.7 displays the marginal average impact of school practices on math scores, with practices in bold being positive and significant. As suspected, the items which are directly referring to the teacher doing extra effort for the students are negatively correlated with performance, while other items display positive correlations. This indicates that the set of items chosen are entangled measurements of at least two distinct phenomena First, an indication of the proficiency of students as lower achievers which requires more effort from the teacher and second, the actual measurement of the teacher's pedagogy.

Table 1. 8. Marginal impact of school practices on average math score, 2003-2012 pooled PISA data for Spain

Name	Variable	Mean estimate	T-stat	Significant at 5% level
	The teacher...			
TEACHSUP1	**shows an interest in every student's learning**	**6.6**	**7.6**	**YES**
TEACHSUP2	gives extra help when students need it	1.2	1.6	NO
TEACHSUP3	helps students with their learning	-1.4	1.6	NO
TEACHSUP4	continues teaching until the students understand	-0.3	0.4	NO
TEACHSUP5	**gives students an opportunity to express opinions**	**3.7**	**4.9**	**YES**
STUREL1	Students get along well with most teachers	-0.2	0.2	NO
STUREL2	**Most teachers are interested in students' well-being**	**4.6**	**4.5**	**YES**
STUREL3	**Most of my teachers really listen to what I have to say**	**5.1**	**5.1**	**YES**
STUREL4	If I need extra help, I will receive it from my teachers	-0.3	0.3	NO
STUREL5	Most of my teachers treat me fairly	-2.3	2.4	YES

Source: Author's calculations. Controls: other indexes of school practices. Standard deviations are computed by clustering observations by schools. Number of observations: 17,379. Adjusted R^2 is 0.32. OLS regressions.

For instance, the fact that a teacher is showing interest in every student's learning is a positive feature and should not a priori be linked to the overall level of performance of the student. The same applies to the teacher giving to student the opportunity to express their opinion.

Table 1.9 New variables of Teacher's practices

	Unbiased items	Biased items
New variable	TEACHPRAC	TEACHBIAS
List of variables	TEACHSUP, TEACHSUP5, STUREL2, STUREL3	TEACHSUP2-TEACHSUP4, STUREL1, STUREL4, STUREL5
KMO	0.65	0.78

On the opposite, the remaining items are likely to suffer from endogeneity bias and are therefore not well identified. For instance, the question "If I need extra help, I will receive it from my teachers" is clearly biased toward low-achievers as top performers are less likely to need extra help in any case.

To compile less biased measures of teacher practices, we now regroup the ten items used to initially measure the support from teachers and the student-teacher relationships differently, grouping first the measures that are unlikely to be biased and second the measures that may be biased. We then perform a factor analysis of the two groups to obtain two different new indexes. The list of variables in the new groups and the Kaiser-Meyer-Olkin (KMO) statistics, which indicate whether the items are enough correlated for the factor analysis to be relevant, are displayed in the Table 1.9 below. KMO tests suggest that the factor analysis is indeed valid.

When introduced along with the other school practices variables in an OLS regression with the average math score as a dependent variable, the new variable TEACHPRAC is not only positively correlated with student's performance in math but also highly significant (T-stat is 12.1), while the other variable TEACHBIAS is not correlated at all with student's performance.

Doing an analysis at the item level of the components of the index "belong at school" also reveals (Table 1.11) that the items may be measuring two distinct things: first the ability to have many friends and second, the overall feeling of fitting socially in the school. A factor analysis of the components of the "BELONG" index indeed shows that they can be decomposed into two distinct factors.

Table 1. 10. Marginal impact of "belonging to school" items on math performance and factor decomposition

	Items	Mean	Standard Error	T-Stat	MAKFRND	LIKSCHL
BELONG1	I feel like an outsider (or left out of things) at school	8.9	1.2	7.4	0.22	0.25
BELONG 2	I make friends easily at school	14.9	1.4	13.1	-0.2	0.32
BELONG3	I feel like I belong at school	-2.8	1.1	2.5	-0.2	0.25
BELONG4	I feel awkward and out of place in my school	4.2	1	4.1	0.22	0.2
BELONG 5	Other students seem to like me	5.8	1.1	5.1	-0.15	0.24
BELONG 6	I feel lonely at school	-3.4	1.3	2.7	0.27	0.26

Source: Author's calculations. Controls: other indexes of school practices. Standard deviations are computed by clustering observations by schools. Number of observations: 16,842. Adjusted R^2 is 0.34. OLS regressions.

The first one, which we label "MAKFRND", is associated with the ability to make friends while the second one, which we labelled "LIKSCHL" (representing the sense of belonging to school) is not. It is to be noted that because of the way item responses are coded in the PISA survey (1 for strongly agree to 4 for strongly disagree) item answers which are negatively correlated with a variable actually traduce that students who agree with the item assertion tend to be display a higher value of the correlated variable. Here, students who "make friends easily at school" (thus having lower values for the item BELONG2) display a higher MAKFRND index but a lower "LIKSCHL" (see Table 1.10). Similarly, students reporting that they are appreciated by their peers (item BELONG5) display a higher MAKFRND index but a lower LIKSCHL index and the same applies for item BELONG3. By construction using the factor analysis, the index LIKSCHL is measuring to which extent students feel" right" in their school environment, regardless their individual ability to make friends. This latter index is positively correlated with math performance while the first one is not. This negative correlation between the ability to make friends and math performance suggests some tension between certain interpersonal skills and math performance.

A similar type of item-level analysis of the component of "being interested in math" also leads to contrasted results (Table 1.11). However, the number of items is too limited to be able to disentangle two clear distinct factors. Pupils interested mathematics display higher PISA scores while pupils declaring to be enjoying mathematics display lower scores. We then build a new index MATENJ based only on items 2 and 4.

Table 1. 11. Marginal impact of "interested in math" items on math performance

Name	Items	Mean	T-Stat
INTMAT1	I enjoy reading about mathematics	-3.4	3.7
INTMAT2	I look forward to my mathematics lessons	12.8	12.3
INTMAT3	I do mathematics because I enjoy it	-5.6	5.1
INTMAT4	I am interested in the things I learn in mathematics	8.6	8.7

Source: Author's calculations. Controls: other indexes of school practices. Standard deviations are computed by clustering observations by schools. Number of observations: 16,842. Adjusted R^2 is 0.34. OLS regressions.

Having now 12 indexes measuring overall school practices, out of which seven are likely to be exogenous, we now look at missing values. The number of missing values for the items used to build these indexes appears to be quite large. However, one can exploit the correlations between items of each group to predict the value of the principal factors. To do so, the principal factor is regressed on all but the missing items, allowing to project values for the index when only one underlying variable is missing. Furthermore, we compute the weighted average of each indicator at the school level. We also use the correlation between the individual values and the school average to fill in missing datA The principal factor is regressed over the school average. It is to be noted that the use of raw and imputed data give similar results when computing the marginal impacts of school practices on math scores (see Table 1.12).

We now ask ourselves if the group of indicators is actually measuring the expected effects by regressing the average math performance over them all, at both the individual and the school level. Doing so, all the unbiased indicators are now properly correlated with math performance, statistically significant and the signs of both the marginal individual impact and the marginal school impact are consistent. This indicates that what we are now measuring does not traduce gross endogeneity biases at the individual but rather measures something meaningful about actual school practices. There is, however, an exception as pupils who tend to self-evaluate their math proficiency higher are in fact less performing students.

Table 1.12 Marginal impact of school practices on average math score, 2003-2012 pooled PISA data for Spain

School practices	Level	Mean	T-stat	Exogeneity
Math anxiety	Individual	-19.9	4.4	YES
Math anxiety	School average	-8.1	12.7	YES
Math self-concept	Individual	-14.5	3.5	NO
Math self-concept	School average	18.5	24.7	YES
Disciplinary climate	Individual	6.7	2.8	YES
Disciplinary climate	School average	7.7	12.9	YES
Interest for math	Individual	-63.1	6	NO
Interest for math	School average	-23.8	16.5	NO
Intrinsic motivation	Individual	6.2	1.4	YES
Intrinsic motivation	School average	7.2	12.7	YES
Attitude toward school	Individual	8.9	2.7	YES
Attitude toward school	School average	4.4	8.7	YES
Math Self-Efficacy	Individual	38.6	12	YES
Math Self-Efficacy	School average	32.2	54	YES
TEACHPRAC	Individual	15.2	2.8	YES
TEACHPRAC	School average	10	11.5	YES
TEACHBIAS	Individual	1.2	0.3	NO
TEACHBIAS	School average	-1.3	1.6	NO
MAKFRND	Individual	9.6	3	YES
MAKFRND	School average	-1.7	3	NO
LIKSCHL	Individual	14.2	3.2	YES
LIKSCHL	School average	11.4	16	YES
MATENJ	Individual	45.5	5	YES
MATENJ	School average	14.5	12	YES

Source: Author's calculations. No other controls. Standard deviations are computed by clustering observations by schools. Number of observations: 35,785. Adjusted R^2 is 0.32. OLS regressions.

Decomposing changes in math teaching practices between 2003 and 2012

PISA scores in math have not changed at all between 2003 and 2012, with the average score in PISA 2003 being 485 points in the OECD scale and 484 in 2012. Nevertheless, this section conducts a counterfactual analysis of the evolution of math practices and identifies the overall impact of changes in math attitudes, motivations, and practices on math scores between 2003 and 2012. We extract the most likely exogenous variables from Table 1.13 (both at individual and school average) and rearrange them into three groups: math attitudes, school climate, and teacher practices at school.

Table 1.13 School and Teacher Practices by groups of variables

Group of Variables	Variables Included (Individual and School Level)	Acronym
Math Motivations	Math Anxiety	ANXMAT
Math Motivations	Math Enjoying	MATENJ
Math Motivations	Intrinsic Motivation on Math	INSTMAT
School Climate	Disciplinary Climate	DISCLIMA
School Climate	Attitude Towards School	ATSCHL
School Climate	School Belonging	LIKSCHL
Teacher Practices	Math Self-Efficacy	MATHEFF
Teacher Practices	Teacher Practices and Relation with Students	TEACHPRAC

Source: PISA 2003 and 2012.

These three groups of variables are likely to be correlated both with individual and family factors but also school practices. However, for policy purposes, we assume that it is more likely that math motivations' variables may be more related to the student characteristics, while school climate and teaching practices are more in hands of policy-makers. In the resulting production function already described:

$$Y_i = \alpha + X_i'\beta + P_i'\mu + Z_i'\delta + M_i'\eta + \epsilon_i$$

(6)

the variable M_i distinguishes the factors related to Math Motivations, School Climate, and Teacher Practices (Table 1.13). Once the RIF-regressions are estimated, we decompose score differences between 2003 and 2012 by groups of variables for different percentiles of the distribution following the RIF-regression approach.

As in the previous section, the full set of characteristics from equation (6) are displayed Table 1.14 for years 2003 and 2012. In this table, we can see the evolution of the student and school covariates from 2003 to 2012. Changes of variables are many. At the individual level, there are changes for immigration variables (more first generation immigrants), an increase of mother participation on labor market from 2003 to 2009 (with full-time job) and decrease of fathers that work full-time (through crisis effect, from 2009 to 2012), an overall improvement of socio-economic conditions. At the school level, increased teacher-student ratio confirms a decrease in teacher shortages. Regarding indexes of attitudes, motivations, discipline, and math practices, changes are mixed: while the disciplinary climate and the attitudes towards school slightly improved, the students feel less at ease at school. Finally, student reported self-efficacy improved, although math teaching practices worsened.

The impact of the variables displayed in Table 1.A1.5 on student scores is described in the Annex, which shows linear regressions for the math scale in both years 2003 and 2012, following model (6). Results display the impact of repetition, gender, immigration status, pre-school education, family structure (including students unaware of their family structure in 2012), and socio-economic status. All the indexes defined to capture teachers' practices and school learning environment happen to have a significant impact on quality, although individual and school effects tend to outbalance each other for some cases, such as for the Intrinsic Motivation in Math.

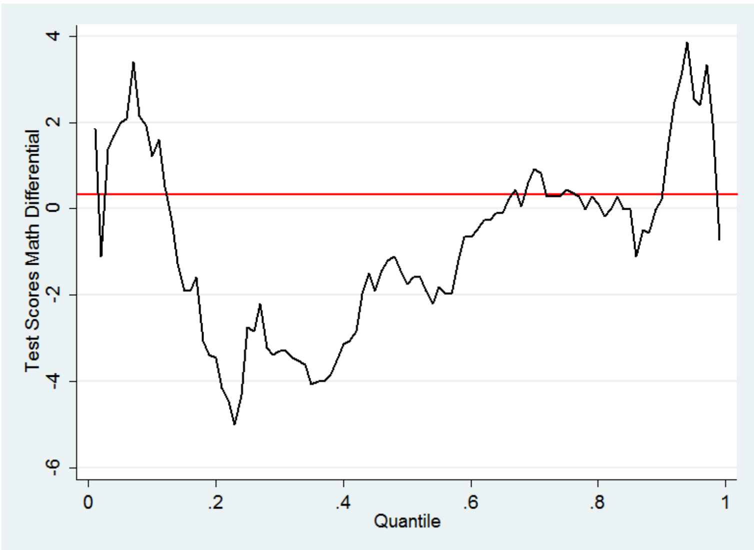
Table 1.14 Student and School Characteristics, 2003-2012

Variables	2003			2012		
	Obs	Mean		Obs	Mean	
Individual			***Individual***			***Individual***
Age	10791	15.85	Age	10791	15.85	Age
Grade	10791	9.67	Grade	10791	9.67	Grade
Girl	10791	0.51	Girl	10791	0.51	Girl
Foreign Language	10751	0.16	Foreign Language	10751	0.16	Foreign Language
Parental Highest Education (Years)	10271	11.11	Parental Highest Education (Years)	10271	11.11	Parental Highest Education (Years)
Immigrant First Generation	10632	0.03	Immigrant First Generation	10632	0.03	Immigrant First Generation
Pre-School 2+ Years	10643	0.84	Pre-School 2+ Years	10643	0.84	Pre-School 2+ Years
Single Parent Family	10791	0.14	Single Parent Family	10791	0.14	Single Parent Family
Family Missing Information	10791	0.01	Family Missing Information	10791	0.01	Family Missing Information
Father Full Time Employed	10791	0.79	Father Full Time Employed	10791	0.79	Father Full Time Employed
Mother Full Time Employed	10791	0.39	Mother Full Time Employed	10791	0.39	Mother Full Time Employed
Index of Economic, Social and Cultural Status	10687	-0.52	Index of Economic, Social and Cultural Status	10687	-0.52	Index of Economic, Social and Cultural Status
Peer Characteristics (School)			***Peer Characteristics (School)***			***Peer Characteristics (School)***
Index of Economic, Social and Cultural Status (School)	10791	-0.52	Index of Economic, Social and Cultural Status (School)	10791	-0.52	Index of Economic, Social and Cultural Status (School)
Share of Immigrants First Generation (School)	10791	0.03	Share of Immigrants First Generation (School)	10791	0.03	Share of Immigrants First Generation (School)
School Characteristics			***School Characteristics***			***School Characteristics***
Teacher Shortage	10401	-0.46	Teacher Shortage	10401	-0.46	Teacher Shortage
Rural	10791	0.28	Rural	10791	0.28	Rural
Public	10791	0.64	Public	10791	0.64	Public
Math Attitudes and Motivations			***Math Attitudes and Motivations***			***Math Attitudes and Motivations***
Math Anxiety	10755	0.21	Math Anxiety	10755	0.21	Math Anxiety
Math Anxiety (School)	10791	0.21	Math Anxiety (School)	10791	0.21	Math Anxiety (School)
Math Enjoying	10760	0.06	Math Enjoying	10760	0.06	Math Enjoying
Math Enjoying (School)	10791	0.06	Math Enjoying (School)	10791	0.06	Math Enjoying (School)
Intrinsic Motivation Math	10753	-0.08	Intrinsic Motivation Math	10753	-0.08	Intrinsic Motivation Math
Intrinsic Motivation Math (School)	10791	-0.08	Intrinsic Motivation Math (School)	10791	-0.08	Intrinsic Motivation Math (School)
School Environment			***School Environment***			***School Environment***
Disciplinary Climate	10755	-0.14	Disciplinary Climate	10755	-0.14	Disciplinary Climate
Disciplinary Climate (Peers)	10791	-0.15	Disciplinary Climate (Peers)	10791	-0.15	Disciplinary Climate (Peers)

Attitude Towards School	10755	0.17	Attitude Towards School	10755	0.17	Attitude Towards School
Attitude Towards School (School)	10791	0.17	Attitude Towards School (School)	10791	0.17	Attitude Towards School (School)
School Sense of Belonging	10760	0.06	School Sense of Belonging	10760	0.06	School Sense of Belonging
School Sense of Belonging (School)	10791	0.06	School Sense of Belonging (School)	10791	0.06	School Sense of Belonging (School)
Teaching Practices			***Teaching Practices***			***Teaching Practices***
Math Self-Efficacy	10752	-0.12	Math Self-Efficacy	10752	-0.12	Math Self-Efficacy
Math Self-Efficacy (School)	10791	-0.12	Math Self-Efficacy (School)	10791	-0.12	Math Self-Efficacy (School)
Teaching Practices	10760	0.12	Teaching Practices	10760	0.12	Teaching Practices
Teaching Practices (School)	10791	0.12	Teaching Practices (School)	10791	0.12	Teaching Practices (School)

Notea: Data from PISA 2003 and 2012.

Figure 1.2 Change in Score by quantiles in math between 2003 and 2012



Source: PISA 2003 and 2012 student scores. Red line represents average change

Figure 1.2 plots differences in math scores across the distribution of achievers between 2003 and 2012. Although the average scores stalled, students' performance evolved unevenly during the last decade, with low and top achievers securing some gains and median ones declining in performance. This pattern is similar to what was observed in the 2009-2012 analysis, although top performers did improve their scores between 2003 and 2012.

Table 1.15 shows the Oaxaca-Blinder decomposition for math, by grouping different variable following equation (6), but a detailed description of the effect of each variable can be seen in the Technical Annex (Table 1.A6). Changes in observables induced an average increase of 7 points, while the negative contribution of the unexplained part suggests a decline in the overall efficiency of the education system between 2003 and 2012. This, together with results in previous section, suggests that unexplained factors are outweighing explanatory effects, no matter the directions in which these go. On average, the counterfactual exercise shows an improvement of individual factors and math practices of students, and a slightly negative effect of math motivations. A more detailed look displays that explained factors have been significantly important for best performing students (through high effect of individual factors –grade- and learning practices) and insignificant, if not negative, for lowest performing students (effect of teaching practices was lower, and the effect of individual characteristics was negative, mainly due to changes in grade composition, economic crisis –from 2008-, and others).

Table 1.15 Math Scale: Decomposition of Change in Scores in 2003-2012 by mean and quantiles

Variables	Mean	P 10	P 30	P 50	P 70	P 90
Explained (1)	7.200*	-3.235	-0.273	9.676**	13.65***	12.94***
	(3.678)	(3.666)	(2.796)	(4.632)	(3.778)	(2.974)
Individual	-3.268**	-6.485***	-5.278***	-3.599*	0.272	1.363
	(1.535)	(1.514)	(1.294)	(1.898)	(1.419)	(1.021)
Individual (School Average)	6.065***	2.590	2.614**	9.012***	7.041***	5.308***
	(1.384)	(1.579)	(1.215)	(1.860)	(1.493)	(1.165)
School	0.162	-0.182	0.129	0.484	0.324	0.0355
	(0.463)	(0.559)	(0.421)	(0.570)	(0.575)	(0.448)
Math Motivations	-3.461***	-3.946***	-1.659*	-3.748***	-2.961***	-2.772***
	(1.003)	(1.429)	(0.878)	(1.314)	(1.081)	(1.015)
School Climate	1.259**	1.119	0.928*	1.770**	0.763	1.089*
	(0.569)	(0.698)	(0.489)	(0.784)	(0.566)	(0.586)
Teaching Practices	6.443***	3.669*	2.993**	5.756***	8.216***	7.914***
	(1.729)	(2.200)	(1.413)	(2.017)	(1.976)	(1.602)
Unexplained (2)	-6.129*	3.171	-2.503	-8.649*	-11.30***	-11.01***
	(3.411)	(6.270)	(4.368)	(4.737)	(3.922)	(2.625)
Difference (1+2)	1.070	-0.0640	-2.776	1.027	2.353	1.930

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1. Variables effects are grouped after a detailed decomposition as shown in Annex.

CONCLUSION

This study analyzes the evolution of PISA results during the last decade. We argue that students of distinct abilities may not be affected homogeneously by changes in determining factors. To exhibit such differences, we use recent statistical techniques, RIF-regressions, from the labor discrimination literature to compute counterfactual quantile decompositions. We address first the changes of socio-economic endowments during the Great Recession in Spain and we find how the crisis had a significant impact on school performance, by affecting father full-time employment -consistently with the literature - and reducing home possessions. Moreover, this impact was not even across students, hitting more median achievers whose fathers lost their jobs. The fact that students' performance remained fairly stable or slightly improved in spite of a harsh crisis which actually did negatively affect performance, allow us to advance that it may be possible that Spain would have achieved significant gains in PISA scores in the absence of the crisis. Although the actual causes of the education system overall improvement cannot be identified, one can furthermore expect that Spanish students' performance may continue to improve in the medium-run, providing a durable real economic recovery.

The study extends the use of the RIF-regressions to analyze the respective impact of students attitudes and motivations, school environment and teacher practices on math performance. We argue that methodological biases of students responses can be neutralized. We then conduct a specific analysis of variables related to math motivations, attitudes, enjoyment of math, as well as variables related to discipline in school, attitudes, and sense of belonging. Finally we construct exogenous measures of teacher practices and their relations to students. This exercise of variable description is valuable in itself as it eases a better understanding of these measurements and draws options for future changes. Decomposing the differences of scores show that the overall school environment indeed changed in the last decade and that these evolutions had a significant impact on students' performance. In particular, we find that on average teachers' practices contributed negatively to math performance although student strategies had a positive impact. Moreover, results show that students are less motivated than they used to be. The exercise also suggests that the overall efficiency of the education system declined between 2003 and 2009. A more detailed look displays that changes in observable factors had significant positive impact on the best performing students (through high effect of individual factors – grade- and teaching practices) and insignificant, if not negative, for the lowest achievers.

This type of analysis contributes to the education policy analysis by bringing a modern technique for decomposing scores of two different populations at different points of each distribution. The use of decomposition techniques on linear regressions offers a general picture on the average students. In this paper, we argue that it is crucial for policy-makers to understand the trade-offs of potential education policies: some may be successful for top performing students, but may be harming students at the bottom of the scale. There are relevant policy questions for Spain that could be analyzed in the future with the methodology proposed by Firpo, Fortin and Lemieux, such as the worrying structure of repetition in Spain and its effect on performance, the differences between public and private schools and its efficiency, the differences between immigrants and native population, or inter-regional differences in performance.

ANNEX

Table 1. A1 Linear Regressions of scores by each scale and year

7	Math		Read		Science	
	2009	2012	2009		2009	2012
Individual Non-Crisis				***Individual Non-Crisis***		
Age	-2.161	0.187	5.397**	Age	-2.161	0.187
	(2.503)	(2.289)	(2.412)		(2.503)	(2.289)
Grade	63.97***	60.75***	57.56***	Grade	63.97***	60.75***
	(1.319)	(1.587)	(1.237)		(1.319)	(1.587)
Girl	-32.09***	-25.77***	15.96***	Girl	-32.09***	-25.77***
	(1.448)	(1.357)	(1.350)		(1.448)	(1.357)
Foreign Language	-1.167	-2.612	-5.138**	Foreign Language	-1.167	-2.612
	(3.645)	(2.448)	(2.489)		(3.645)	(2.448)
Parental Highest Education (Years)	0.197	0.514**	0.182	Parental Highest Education (Years)	0.197	0.514**
	(0.275)	(0.236)	(0.240)		(0.275)	(0.236)
Immigrant First Generation	-18.29***	-11.72***	-19.85***	Immigrant First Generation	-18.29***	-11.72***
	(3.212)	(2.800)	(2.662)		(3.212)	(2.800)
Immigrant Second Generation	-14.27**	-9.916**	-8.413	Immigrant Second Generation	-14.27**	-9.916**
	(6.983)	(4.977)	(6.163)		(6.983)	(4.977)
Pre-School 1 year	2.499	8.306**	4.691	Pre-School 1 year	2.499	8.306**
	(4.402)	(3.658)	(4.880)		(4.402)	(3.658)
Pre-School 2+ Years	16.13***	20.38***	13.82***	Pre-School 2+ Years	16.13***	20.38***
	(3.924)	(2.884)	(4.077)		(3.924)	(2.884)
Single Parent Family	-1.034	6.010**	0.785	Single Parent Family	-1.034	6.010**
	(2.281)	(2.427)	(2.251)		(2.281)	(2.427)
Family Missing Information	-36.21***	-13.85***	-32.38***	Family Missing Information	-36.21***	-13.85***
	(7.919)	(3.149)	(8.054)		(7.919)	(3.149)
Individual Non-Crisis (School Average)				***Individual Non-Crisis (School Average)***		
Immigrant Second Generation (School)	72.00	-21.21	109.2*	Immigrant Second Generation (School)	72.00	-21.21
	(62.54)	(63.00)	(55.76)		(62.54)	(63.00)
Immigrant First Generation (School)	6.396	30.17**	23.25	Immigrant First Generation (School)	6.396	30.17**
	(14.70)	(14.18)	(14.98)		(14.70)	(14.18)
Pre-School 2+ Years (School)	-6.312	11.77	28.30**	Pre-School 2+ Years (School)	-6.312	11.77
	(14.80)	(12.79)	(13.44)		(14.80)	(12.79)
Foreign Language (School)	4.544	-8.411	0.548	Foreign Language (School)	4.544	-8.411
	(7.547)	(7.331)	(7.980)		(7.547)	(7.331)
Percentage of Girls in School	60.40***	-8.704	49.66	Percentage of Girls in School	60.40***	-8.704
	(20.41)	(11.26)	(33.07)		(20.41)	(11.26)
Parental Highest Education Years (School)	4.129**	1.959	3.077*	Parental Highest Education Years (School)	4.129**	1.959
	(1.713)	(1.535)	(1.585)		(1.713)	(1.535)
Average Grade in School	-5.780	7.593	4.922	Average Grade in School	-5.780	7.593

	(7.518)	(6.465)	(7.676)		(7.518)	(6.465)
Individual Crisis				***Individual Crisis***		
Father Full Time Employed	2.832*	2.628*	0.866	Father Full Time Employed	2.832*	2.628*
	(1.534)	(1.499)	(1.541)		(1.534)	(1.499)
Mother Full Time Employed	1.998	0.392	-0.364	Mother Full Time Employed	1.998	0.392
	(1.290)	(1.372)	(1.270)		(1.290)	(1.372)
Index of Highest Parental Occupation	0.411***	0.378***	0.408***	Index of Highest Parental Occupation	0.411***	0.378***
	(0.0504)	(0.0448)	(0.0527)		(0.0504)	(0.0448)
Index of Home Educational Resources	3.035***	1.906**	2.195**	Index of Home Educational Resources	3.035***	1.906**
	(0.915)	(0.883)	(0.874)		(0.915)	(0.883)
Index of Cultural Possessions	9.943***	10.67***	10.87***	Index of Cultural Possessions	9.943***	10.67***
	(0.875)	(0.960)	(0.793)		(0.875)	(0.960)
Index of Wealth	-2.292**	-2.029*	-4.276***	Index of Wealth	-2.292**	-2.029*
	(1.059)	(1.222)	(1.056)		(1.059)	(1.222)
Individual Crisis (School Average)				***Individual Crisis (School Average)***		
Father Full Time Employed (School)	55.09***	51.77***	7.333	Father Full Time Employed (School)	55.09***	51.77***
	(14.45)	(12.97)	(14.40)		(14.45)	(12.97)
Mother Full Time Employed (School)	14.63	9.069	34.64***	Mother Full Time Employed (School)	14.63	9.069
	(12.44)	(10.81)	(11.80)		(12.44)	(10.81)
Index of Highest Parental Occupation (School)	-0.285	-0.112	-0.00753	Index of Highest Parental Occupation (School)	-0.285	-0.112
	(0.382)	(0.276)	(0.377)		(0.382)	(0.276)
Index of Home Educational Resources (School)	*1.666*	*1.971*	*5.524*	*Index of Home Educational Resources (School)*	*1.666*	*1.971*
	(6.493)	(5.551)	(6.401)		(6.493)	(5.551)
Index of Cultural Possessions (School)	28.70***	15.32**	18.77**	Index of Cultural Possessions (School)	28.70***	15.32**
	(7.590)	(6.184)	(7.432)		(7.590)	(6.184)
Index of Wealth (School)	-13.93***	-10.82*	-9.152*	Index of Wealth (School)	-13.93***	-10.82*
	(5.259)	(5.665)	(5.143)		(5.259)	(5.665)
School Resources-Crisis				***School Resources-Crisis***		
Quality of School Resources	-1.906	-2.259	-0.690	Quality of School Resources	-1.906	-2.259
	(2.973)	(2.039)	(2.548)		(2.973)	(2.039)
Shortage of subject teacher	2.444	-2.702	0.927	Shortage of subject teacher	2.444	-2.702
	(3.038)	*(2.993)*	*(3.416)*		*(3.038)*	*(2.993)*
Shortage of Library Materials	-4.449*	-2.916	-1.799	Shortage of Library Materials	-4.449*	-2.916
	(2.390)	(1.964)	(2.301)		(2.390)	(1.964)
Shortage of Instruction Materials	-1.571	-2.012	0.374	Shortage of Instruction Materials	-1.571	-2.012
	(2.362)	(2.362)	(2.108)		(2.362)	(2.362)
School Factors				***School Factors***		

1. Determinants of Changes in Education Quality In Spain

Rural	7.239**	1.693	-0.140	Rural	7.239**	1.693
	(3.308)	(2.949)	(3.399)		(3.308)	(2.949)
Public School	5.939*	-0.520	1.533	Public School	5.939*	-0.520
	(3.414)	(3.088)	(3.366)		(3.414)	(3.088)
Index of Disciplinary Climate at School	10.20**	3.411	8.748*	Index of Disciplinary Climate at School	10.20**	3.411
	(4.545)	(3.062)	(4.476)		(4.545)	(3.062)
Constant	175.9***	158.2***	21.88	Constant	175.9***	158.2***
	(51.16)	(46.31)	(54.59)		(51.16)	(46.31)
Observations	22,316	21,749	22,283	Observations	22,316	21,749
R-squared	0.461	0.451	0.454	R-squared	0.461	0.451

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1.

Table 1. A2 Math Scale: Detailed decomposition of Change in Scores in 2009-2012 by mean and percentiles

Variables	Mean	P 10	P 30	P 50	P 70	P 90
Individual Non-Crisis						
Age	0.000832	-0.0178	-0.00528	-0.00330	0.0170	0.00463
	(0.0102)	(0.0276)	(0.0201)	(0.0177)	(0.0312)	(0.0176)
Grade	2.205*	2.140*	3.937*	2.694*	2.648*	0.898*
	(1.184)	(1.156)	(2.115)	(1.448)	(1.425)	(0.485)
Girl	0.221	0.0976	0.226	0.233	0.396	0.258
	(0.241)	(0.108)	(0.247)	(0.255)	(0.432)	(0.282)
Foreign Language	-0.0152	-0.00188	-0.00253	-0.0231	-0.0298	-0.0345
	(0.0576)	(0.0259)	(0.0276)	(0.0901)	(0.114)	(0.130)
Parental Highest Education (Years)	-0.384**	-0.162	-0.332	-0.563*	-0.479	-0.369
	(0.193)	(0.337)	(0.360)	(0.295)	(0.375)	(0.271)
Immigrant First Generation	-0.0228	-0.0129	-0.0375	-0.0324	-0.0375	-0.0186
	(0.0936)	(0.0537)	(0.154)	(0.133)	(0.154)	(0.0763)
Immigrant Second Generation	-0.0427	-0.0105	-0.0104	-0.0848	-0.130	-0.0347
	(0.0306)	(0.0579)	(0.0500)	(0.0582)	(0.0799)	(0.0338)
Pre-School 1 year	-0.0367	-0.0907	-0.0781	-0.00708	-0.0181	0.0353
	(0.0481)	(0.119)	(0.106)	(0.0287)	(0.0399)	(0.0473)
Pre-School 2+ Years	-0.144	-0.216	-0.197	-0.103	-0.166	-0.0261
	(0.151)	(0.231)	(0.212)	(0.112)	(0.178)	(0.0360)
Single Parent Family	-0.224**	-0.399***	-0.423*	-0.174	-0.256	0.127
	(0.0966)	(0.149)	(0.221)	(0.141)	(0.195)	(0.134)
Family Missing Information	-0.708***	-0.856***	-1.334***	-0.748***	-0.726**	-0.255
	(0.166)	(0.298)	(0.350)	(0.270)	(0.342)	(0.218)
Individual Non-Crisis (School Average)						
Immigrant Second Generation (School)	-0.0720	-0.0355	-0.0838	0.00571	-0.0458	-0.0285
	(0.216)	(0.306)	(0.345)	(0.235)	(0.329)	(0.169)
Immigrant First Generation (School)	0.0276	0.00246	0.0298	0.0357	0.0698	0.0278
	(0.236)	(0.0270)	(0.256)	(0.306)	(0.598)	(0.238)
Pre-School 2+ Years (School)	-0.134	0.145	-0.0470	-0.222	-0.242	-0.294
	(0.167)	(0.240)	(0.226)	(0.219)	(0.308)	(0.236)
Foreign Language (School)	-0.0541	-0.0799	-0.0583	-0.0490	-0.0713	-0.0484
	(0.186)	(0.273)	(0.207)	(0.174)	(0.251)	(0.168)
Percentage of Girls in School	0.277	-0.0440	0.183	0.183	0.860	0.225
	(0.364)	(0.475)	(0.514)	(0.462)	(0.685)	(0.411)
Parental Highest Education Years (School)	-2.155	0.0605	-2.056	-2.324	-4.211	-2.592
	(1.702)	(1.766)	(2.912)	(2.265)	(2.995)	(1.597)
Average Grade in School	0.197	0.239	0.183	0.170	0.226	0.0830
	(0.226)	(0.284)	(0.330)	(0.260)	(0.345)	(0.191)
Individual Crisis						
Father Full Time Employed	-0.141*	-0.145	-0.119	-0.113	-0.379*	-0.178
	(0.0837)	(0.155)	(0.184)	(0.138)	(0.198)	(0.129)
Mother Full Time Employed	0.00100	0.000537	0.00323	0.00213	0.0109	6.42e-05
	(0.00604)	(0.00571)	(0.0175)	(0.0123)	(0.0544)	(0.00650)
Index of Highest Parental Occupation	0.467	-0.0201	0.390	0.625	1.073	0.571
	(0.319)	(0.0873)	(0.284)	(0.431)	(0.730)	(0.392)
Index of Home Educational Resources	-0.130*	-0.0584	-0.0576	-0.125	-0.317*	-0.231**
	(0.0706)	(0.103)	(0.121)	(0.103)	(0.166)	(0.105)
Index of Cultural Possessions	-0.795***	-0.701**	-1.139***	-0.884***	-1.095***	-0.560***
	(0.286)	(0.273)	(0.428)	(0.330)	(0.411)	(0.217)
Index of Wealth	-0.142	0.218	-0.306	-0.307*	-0.434*	-0.213
	(0.0979)	(0.169)	(0.215)	(0.167)	(0.229)	(0.140)
Individual Crisis (School Average)						
Father Full Time Employed (School)	-2.881***	-2.302**	-4.056***	-3.811***	-3.842**	-2.289***
	(0.874)	(0.901)	(1.343)	(1.167)	(1.537)	(0.829)
Mother Full Time Employed (School)	0.0320	0.0772	0.0507	0.0382	0.0171	0.0207

Variables	Mean	P 10	P 30	P 50	P 70	P 90
	(0.118)	(0.274)	(0.189)	(0.141)	(0.0965)	(0.0850)
Index of Highest Parental Occupation (School)	-0.0470	-0.153	-0.243	0.0593	0.0283	0.0228
	(0.148)	(0.329)	(0.519)	(0.193)	(0.208)	(0.130)
Index of Home Educational Resources (School)	-0.138	-0.330	0.169	0.324	-0.248	-0.165
	(0.386)	(0.477)	(0.690)	(0.511)	(0.715)	(0.417)
Index of Cultural Possessions (School)	-1.215*	-0.325	-2.362**	-1.630*	-2.124*	-0.841
	(0.629)	(0.477)	(1.121)	(0.841)	(1.187)	(0.646)
Index of Wealth (School)	-0.674	0.120	-0.485	-0.857	-0.821	-1.110*
	(0.435)	(0.432)	(0.594)	(0.559)	(0.699)	(0.598)
School Resouces-Crisis						
Quality of School Resources	-0.00515	-0.00417	-0.00970	-0.00446	-0.00720	-0.00341
	(0.154)	(0.125)	(0.290)	(0.133)	(0.215)	(0.102)
Math Teacher Shortage	-0.168	0.0229	-0.170	-0.339	-0.295	-0.135
	(0.197)	(0.212)	(0.319)	(0.260)	(0.417)	(0.231)
Shortage of Library Materials	0.594	0.845	0.823	0.672	0.908	0.427
	(0.447)	(0.534)	(0.707)	(0.569)	(0.782)	(0.409)
Shortage of Instruction Materials	-0.00321	0.00515	-0.00641	-0.00336	-0.00496	-0.00199
	(0.108)	(0.173)	(0.216)	(0.113)	(0.167)	(0.0669)
School Factors						
Rural	-0.0350	0.00828	-0.0201	-0.0339	-0.0942	0.0188
	(0.0873)	(0.0760)	(0.119)	(0.101)	(0.201)	(0.0654)
Public School	-0.00830	-0.0301	-0.0656	0.00375	0.0349	0.0331
	(0.0527)	(0.0928)	(0.176)	(0.0637)	(0.119)	(0.0938)
Index of Disciplinary Climate at School	-0.416	-0.529	-0.492	-0.532	-0.679	-0.778*
	(0.385)	(0.402)	(0.601)	(0.545)	(0.708)	(0.464)
Year 2012 Estimated Score	491.4***	380.5***	449.3***	493.9***	541.7***	597.2***
	(2.235)	(2.044)	(3.311)	(2.818)	(3.595)	(1.862)
Year 2009 Estimate Score	488.2***	372.5***	446.5***	491.3***	535.6***	593.3***
	(2.319)	(1.941)	(2.255)	(1.820)	(2.570)	(1.641)
Difference	3.262	8.035***	2.839	2.539	6.095	3.901
	(3.221)	(2.818)	(4.006)	(3.355)	(4.420)	(2.482)
Explained Factors	-6.768**	-2.542	-8.201	-7.927*	-10.46**	-7.454***
	(3.273)	(2.930)	(5.021)	(4.145)	(5.273)	(2.721)
Unexplained Factors	10.03***	10.58***	11.04***	10.47***	16.56***	11.36***
	(2.813)	(2.804)	(4.202)	(3.419)	(4.757)	(2.787)
Constant	-17.63	29.48	-250.2**	-194.2**	-204.9*	55.01
	(68.97)	(113.0)	(106.5)	(86.09)	(115.1)	(87.98)
Observations	44,065	44,065	44,065	44,065	44,065	44,065
Variables	**Mean**	**P 10**	**P 30**	**P 50**	**P 70**	**P 90**

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1.

Table 1. A3 Reading Scale: Detailed decomposition of Change in Scores in 2009-2012 by mean and percentiles

Variables	Mean	P 10	P 30	P 50	P 70	P 90
Individual Non-Crisis						
Age	0.00125	0.00107	0.00804	0.0221	0.0110	0.00563
	(0.0103)	(0.0177)	(0.0218)	(0.0286)	(0.0194)	(0.0137)
Grade	2.214*	1.816*	3.196*	2.017*	1.473*	0.516*
	(1.184)	(0.977)	(1.710)	(1.080)	(0.789)	(0.278)
Girl	0.221	-0.119	-0.196	-0.138	-0.112	-0.0545
	(0.241)	(0.131)	(0.215)	(0.151)	(0.123)	(0.0613)
Foreign Language	-0.0150	-0.0300	-0.0312	-0.00464	-0.0238	-0.0205
	(0.0578)	(0.115)	(0.120)	(0.0248)	(0.0914)	(0.0786)
Parental Highest Education (Years)	-0.382**	0.148	0.235	-0.408*	-0.517**	-0.551**
	(0.192)	(0.316)	(0.306)	(0.230)	(0.263)	(0.219)
Immigrant First Generation	-0.0226	0.00281	-0.0230	-0.0350	-0.0249	-0.0173
	(0.0937)	(0.0159)	(0.0959)	(0.145)	(0.103)	(0.0718)
Immigrant Second Generation	-0.0423	-0.0837	-0.182*	-0.113*	-0.0860*	-0.0937*
	(0.0304)	(0.0651)	(0.106)	(0.0669)	(0.0519)	(0.0515)
Pre-School 1 year	-0.0371	-0.127	-0.0843	-0.0103	-0.0275	-0.0103
	(0.0485)	(0.160)	(0.110)	(0.0260)	(0.0387)	(0.0193)
Pre-School 2+ Years	-0.143	-0.220	-0.151	-0.0440	-0.0613	-0.0349
	(0.152)	(0.235)	(0.165)	(0.0538)	(0.0689)	(0.0401)
Single Parent Family	-0.224**	-0.243	-0.698***	-0.178	0.00376	0.0272
	(0.0963)	(0.169)	(0.214)	(0.131)	(0.133)	(0.104)
Family Missing Information	-0.710***	-0.106	-0.324	-0.230	0.122	-0.118
	(0.166)	(0.303)	(0.322)	(0.200)	(0.187)	(0.176)
Individual Non-Crisis (School Average)						
Immigrant Second Generation (School)	-0.0619	0.331	0.325	0.366	0.226	0.308
	(0.214)	(0.317)	(0.373)	(0.313)	(0.240)	(0.223)
Immigrant First Generation (School)	0.0263	0.0267	0.0528	0.0560	0.0364	0.0221

	(0.230)	(0.234)	(0.462)	(0.490)	(0.319)	(0.193)
Pre-School 2+ Years (School)	-0.137	0.0401	0.102	-0.0965	-0.205	-0.161
	(0.168)	(0.190)	(0.220)	(0.160)	(0.208)	(0.163)
Foreign Language (School)	-0.0548	-0.0226	-0.0427	-0.0215	0.00336	0.0401
	(0.190)	(0.0913)	(0.170)	(0.0957)	(0.0628)	(0.141)
Percentage of Girls in School	0.227	-0.135	-1.272	-0.609	-0.511	-0.344
	(0.356)	(0.306)	(0.839)	(0.550)	(0.534)	(0.327)
Parental Highest Education Years (School)	-2.179	1.094	-1.134	0.146	-1.145	-2.251*
	(1.698)	(1.879)	(3.144)	(2.054)	(1.734)	(1.339)
Average Grade in School	0.177	0.248	-0.0437	0.166	0.133	0.190
	(0.218)	(0.302)	(0.307)	(0.236)	(0.214)	(0.202)
Individual Crisis						
Father Full Time Employed	-0.142*	-0.246	-0.420***	-0.227*	-0.266**	-0.0553
	(0.0836)	(0.162)	(0.162)	(0.125)	(0.127)	(0.0877)
Mother Full Time Employed	0.00106	0.00820	0.00267	0.00207	0.00256	0.000200
	(0.00618)	(0.0393)	(0.0143)	(0.0109)	(0.0132)	(0.00464)
Index of Highest Parental Occupation	0.469	0.0264	0.328	0.381	0.527	0.490
	(0.319)	(0.0850)	(0.242)	(0.266)	(0.362)	(0.337)
Index of Home Educational Resources	-0.130*	-0.202*	-0.434***	-0.297**	-0.288**	-0.161**
	(0.0706)	(0.115)	(0.165)	(0.118)	(0.114)	(0.0779)
Index of Cultural Possessions	-0.794***	-0.499**	-1.091***	-0.743***	-0.893***	-0.463***
	(0.286)	(0.207)	(0.401)	(0.275)	(0.322)	(0.177)
Index of Wealth	-0.143	-0.186	-0.516**	-0.535**	-0.544***	-0.361**
	(0.0979)	(0.159)	(0.227)	(0.216)	(0.210)	(0.149)
Individual Crisis (School Average)						
Father Full Time Employed (School)	-2.894***	-0.946	-1.637	-0.439	-1.028	-0.290
	(0.868)	(0.769)	(1.217)	(0.940)	(0.879)	(0.706)
Mother Full Time Employed (School)	0.0346	0.0609	0.0517	0.0141	-0.0290	-0.0382
	(0.124)	(0.212)	(0.189)	(0.0680)	(0.108)	(0.135)
Index of Highest Parental Occupation (School)	-0.0590	-0.0247	0.00151	0.124	0.278	0.0355
	(0.163)	(0.132)	(0.200)	(0.285)	(0.565)	(0.146)
Index of Home Educational Resources (School)	-0.112	0.278	0.543	0.268	0.213	-0.0182
	(0.383)	(0.496)	(0.623)	(0.444)	(0.466)	(0.357)
Index of Cultural Possessions (School)	-1.299**	-1.056*	-2.006	-1.944**	-0.565	-0.634
	(0.645)	(0.598)	(1.248)	(0.977)	(0.783)	(0.549)
Index of Wealth (School)	-0.654	0.111	0.742	0.389	-0.317	-0.553
	(0.432)	(0.418)	(0.691)	(0.476)	(0.485)	(0.409)
School Resouces-Crisis						
Quality of School Resources	-0.00573	-0.00458	-0.00192	0.00313	-0.00164	0.000615
	(0.145)	(0.116)	(0.0496)	(0.0795)	(0.0420)	(0.0163)
Reading Teacher Shortage	0.0331	0.0314	0.209	0.165	0.172	0.0363
	(0.0876)	(0.0821)	(0.224)	(0.180)	(0.194)	(0.0747)
Shortage of Library Materials	0.604	0.245	0.243	-0.268	0.407	-0.153
	(0.450)	(0.477)	(0.697)	(0.452)	(0.449)	(0.310)
Shortage of Instruction Materials	-0.00354	-0.000229	-0.00465	-0.00566	-0.00301	-0.00421
	(0.124)	(0.00893)	(0.162)	(0.198)	(0.105)	(0.147)
School Factors						
Rural	-0.0328	0.0669	0.0627	0.0263	0.0599	0.0281
	(0.0845)	(0.145)	(0.156)	(0.0866)	(0.129)	(0.0686)
Public School	-0.00999	-0.0319	-0.0433	-0.0335	0.0477	0.0103
	(0.0537)	(0.0958)	(0.133)	(0.103)	(0.134)	(0.0570)
Index of Disciplinary Climate at School	-0.400	0.141	-0.261	-0.0181	-0.362	-0.180
	(0.385)	(0.392)	(0.653)	(0.486)	(0.519)	(0.325)
Year 2012 Estimated Score	491.4***	378.4***	455.3***	500.5***	543.1***	598.1***
	(2.234)	(1.793)	(3.000)	(2.237)	(2.128)	(1.409)
Year 2009 Estimate Score	488.2***	372.2***	449.8***	495.1***	533.4***	586.9***
	(2.321)	(2.012)	(3.462)	(3.605)	(2.528)	(2.236)
Difference	3.270	6.213**	5.456	5.396	9.670***	11.26***
	(3.221)	(2.695)	(4.581)	(4.242)	(3.305)	(2.643)
Explained Factors	-6.679**	0.395	-4.495	-2.250	-3.295	-4.856**
	(3.278)	(2.749)	(4.553)	(3.354)	(2.990)	(2.013)
Unexplained Factors	9.949***	5.819*	9.951**	7.647**	12.96***	16.12***
	(2.810)	(3.011)	(4.809)	(3.762)	(3.097)	(2.697)
Constant	-22.82	165.7	99.94	155.5	138.9	220.0***
	(69.22)	(114.1)	(126.4)	(107.4)	(92.24)	(84.08)
Observations	44,052	44,052	44,052	44,052	44,052	44,052

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1.

Table 1. A4 Science Scale: Detailed decomposition of Change in Scores in 2009-2012 by mean and percentiles

Variables	Mean	P 10	P 30	P 50	P 70	P 90
Individual Non-Crisis						
Age	0.00795	0.000317	0.0120	0.0263	0.00814	-0.0107
	(0.0141)	(0.0115)	(0.0202)	(0.0355)	(0.0234)	(0.0285)
Grade	1.791*	1.283*	2.144*	2.196*	2.098*	1.186*
	(1.005)	(0.724)	(1.203)	(1.233)	(1.179)	(0.670)
Girl	0.147	0.0253	0.124	0.184	0.278	0.285
	(0.152)	(0.0300)	(0.130)	(0.191)	(0.288)	(0.296)
Foreign Language	-0.0399	-0.0363	-0.0425	-0.0317	-0.0392	-0.0158
	(0.145)	(0.133)	(0.156)	(0.117)	(0.145)	(0.0669)
Parental Highest Education (Years)	-0.575**	-0.190	-0.392	-0.701**	-1.093**	-0.958**
	(0.231)	(0.256)	(0.283)	(0.299)	(0.428)	(0.447)
Immigrant First Generation	-0.0284	-0.00771	-0.0433	-0.0383	-0.0329	-0.0405
	(0.104)	(0.0299)	(0.159)	(0.141)	(0.121)	(0.149)
Immigrant Second Generation	-0.0660	-0.0505	-0.0173	-0.0777	-0.124	-0.124*
	(0.0418)	(0.0527)	(0.0441)	(0.0557)	(0.0776)	(0.0753)
Pre-School 1 year	-0.0567	-0.0739	-0.0718	-0.0581	-0.0472	-0.00930
	(0.0793)	(0.104)	(0.102)	(0.0837)	(0.0698)	(0.0276)
Pre-School 2+ Years	-0.152	-0.133	-0.126	-0.162	-0.191	-0.0906
	(0.158)	(0.141)	(0.134)	(0.170)	(0.199)	(0.0994)
Single Parent Family	-0.239**	-0.193*	-0.382**	-0.145	-0.169	-0.0468
	(0.109)	(0.113)	(0.157)	(0.168)	(0.181)	(0.204)
Family Missing Information	-0.587***	-0.447*	-0.648**	-0.854***	-0.213	-0.606**
	(0.172)	(0.238)	(0.266)	(0.235)	(0.349)	(0.283)
Individual Non-Crisis (School Average)						
Immigrant Second Generation (School)	0.106	0.136	0.0896	0.150	0.190	0.228
	(0.224)	(0.209)	(0.260)	(0.271)	(0.325)	(0.313)
Immigrant First Generation (School)	0.0397	0.0143	0.0291	0.0459	0.0776	0.0559
	(0.284)	(0.103)	(0.208)	(0.328)	(0.554)	(0.400)
Pre-School 2+ Years (School)	0.0102	-0.0117	0.0565	0.152	-0.0329	-0.0590
	(0.166)	(0.126)	(0.202)	(0.250)	(0.246)	(0.231)
Foreign Language (School)	-0.0334	0.0104	-0.0392	-0.0258	-0.108	-0.0967
	(0.117)	(0.0478)	(0.140)	(0.106)	(0.362)	(0.324)
Percentage of Girls in School	-0.250	-0.261	-0.191	-0.220	0.0196	-0.423
	(0.308)	(0.214)	(0.432)	(0.428)	(0.420)	(0.437)
Parental Highest Education Years (School)	-1.015	0.525	-0.363	-1.205	-2.566	-3.571
	(2.171)	(1.383)	(2.412)	(3.177)	(3.180)	(2.537)
Average Grade in School	0.0581	0.133	0.0649	-0.174	0.0927	-0.0304
	(0.195)	(0.192)	(0.250)	(0.290)	(0.299)	(0.254)
Individual Crisis						
Father Full Time Employed	-0.230**	-0.0220	-0.232*	-0.215	-0.366*	-0.184
	(0.0979)	(0.114)	(0.130)	(0.151)	(0.197)	(0.195)
Mother Full Time Employed	0.00140	0.00658	-0.000387	0.00249	0.00199	-0.00349
	(0.0104)	(0.0474)	(0.00477)	(0.0185)	(0.0151)	(0.0258)
Index of Highest Parental Occupation	0.414	-0.0291	0.232	0.522	0.844	0.829
	(0.311)	(0.0586)	(0.185)	(0.394)	(0.632)	(0.626)
Index of Home Educational Resources	-0.113*	-0.0592	-0.0566	-0.144	-0.234*	-0.219
	(0.0688)	(0.0808)	(0.0893)	(0.102)	(0.140)	(0.141)
Index of Cultural Possessions	-0.863***	-0.410**	-0.913***	-1.000***	-1.301***	-1.184***
	(0.308)	(0.169)	(0.336)	(0.370)	(0.472)	(0.443)
Index of Wealth	-0.425**	0.0923	-0.459**	-0.652**	-0.876**	-0.687**
	(0.173)	(0.0988)	(0.211)	(0.268)	(0.354)	(0.297)
Individual Crisis (School Average)						
Father Full Time Employed (School)	-1.735*	-0.945	-1.731*	-2.050	-2.654*	-1.839
	(0.948)	(0.628)	(0.995)	(1.269)	(1.537)	(1.284)
Mother Full Time Employed (School)	-0.0184	0.0120	0.0111	-0.0189	-0.0325	-0.116
	(0.0834)	(0.0569)	(0.0642)	(0.0927)	(0.145)	(0.459)
Index of Highest Parental Occupation (School)	-0.0925	-0.0353	-0.129	-0.0409	-0.240	0.0235
	(0.256)	(0.114)	(0.344)	(0.188)	(0.623)	(0.183)
Index of Home Educational Resources (School)	0.114	-0.201	0.751	0.161	0.0490	-0.674
	(0.448)	(0.306)	(0.572)	(0.594)	(0.746)	(0.725)
Index of Cultural Possessions (School)	-1.933**	-1.009**	-1.973**	-2.698**	-2.704**	-1.657*
	(0.789)	(0.511)	(0.858)	(1.102)	(1.129)	(0.906)
Index of Wealth (School)	0.162	-0.371	0.872	0.189	0.402	-0.722
	(0.391)	(0.333)	(0.572)	(0.547)	(0.631)	(0.711)
School Resouces-Crisis						
Quality of School Resources	-0.00206	-0.00300	-0.00129	-0.00371	-0.000861	-0.00153
	(0.0709)	(0.103)	(0.0448)	(0.128)	(0.0308)	(0.0533)
Science Teacher Shortage	-0.0247	-0.102	-0.0700	-0.0123	-0.0158	0.101
	(0.0702)	(0.140)	(0.117)	(0.0748)	(0.0931)	(0.167)

Shortage of Library Materials	0.660	-0.00324	-0.0447	-0.226	-0.122	-0.218
	(0.464)	(0.0954)	(0.181)	(0.293)	(0.241)	(0.300)
Shortage of Instruction Materials	-0.00171	0.673*	0.316	0.637	1.147	1.323*
	(0.0598)	(0.402)	(0.476)	(0.604)	(0.753)	(0.738)
School Factors						
Rural	0.143	0.0514	0.252	0.180	0.207	0.0981
	(0.237)	(0.101)	(0.408)	(0.301)	(0.347)	(0.191)
Public School	-0.0285	-0.0276	-0.0202	-0.0291	-0.0637	0.0275
	(0.0952)	(0.0841)	(0.0818)	(0.107)	(0.201)	(0.113)
Index of Disciplinary Climate at School	-0.368	-0.227	-0.247	-0.309	-0.843	-1.244*
	(0.451)	(0.289)	(0.494)	(0.582)	(0.764)	(0.749)
Year 2012 Estimated Score	504.0***	394.9***	462.8***	507.3***	551.0***	605.9***
	(2.127)	(1.327)	(2.309)	(2.731)	(3.244)	(2.791)
Year 2009 Estimate Score	493.3***	385.9***	452.3***	496.8***	538.4***	594.3***
	(2.376)	(2.538)	(2.655)	(2.454)	(1.816)	(1.995)
Difference	10.77***	9.008***	10.43***	10.55***	12.57***	11.59***
	(3.189)	(2.864)	(3.519)	(3.672)	(3.718)	(3.431)
Explained Factors	-5.329	-1.883	-3.240	-6.653	-8.659*	-10.68**
	(3.379)	(2.044)	(3.701)	(4.605)	(5.160)	(4.145)
Unexplained Factors	16.10***	10.89***	13.67***	17.20***	21.23***	22.26***
	(3.417)	(2.890)	(3.851)	(4.451)	(4.861)	(4.060)
Constant	194.5**	266.8**	81.89	-10.91	-8.668	218.0*
	(78.59)	(115.0)	(100.1)	(97.26)	(104.3)	(120.5)
Observations	43,924	43,924	43,924	43,924	43,924	43,924

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1.

Table 1. A5. Linear Regression in Math Scale, 2003 and 2012

Variables	OLS	
	2003	2012
Individual		
Age	7.833**	-0.841
	(3.038)	(2.093)
Grade	57.82***	53.77***
	(2.102)	(1.353)
Girl	-5.941***	-15.82***
	(2.102)	(1.351)
Foreign Language	-7.721**	-9.730***
	(3.325)	(2.308)
Parental Highest Education (Years)	-0.966**	-0.645**
	(0.429)	(0.260)
Immigrant First Generation	-12.31*	-13.38***
	(7.084)	(2.507)
Pre-School 2+ Years	12.31***	12.50***
	(2.749)	(2.159)
Single Parent Family	-0.282	6.172***
	(2.895)	(2.125)
Family Missing Information	-12.18	-9.644***
	(10.09)	(2.865)
Father Full Time Employed	0.630	3.429**
	(2.408)	(1.386)
Mother Full Time Employed	0.252	1.557
	(1.814)	(1.308)
Index of Economic, Social and Cultural Status	8.823***	10.34***
	(1.791)	(1.088)
Peer Characteristics (School)		
Index of Economic, Social and Cultural Status	17.58***	15.49***
	(4.356)	(3.554)
Share of Immigrants First Generation	60.40***	-8.704
	(20.41)	(11.26)
School Characteristics		
Teacher Shortage	-0.767	-0.965
	(1.519)	(1.985)
Rural	4.797	3.757
	(3.493)	(2.912)
Public	2.135	-0.793
	(4.190)	(3.129)
Math Attitudes and Motivations		
Math Axiety	-14.49***	-11.62***
	(1.283)	(0.991)
Math Anxiety (School)	-7.643	9.916**
	(8.154)	(4.870)
Math Enjoying	6.796***	5.513***
	(1.170)	(1.107)
Math Enjoying (School)	-10.81	-3.198
	(8.536)	(5.539)
Intrinsic Motivation Math	5.249***	5.543***
	(1.133)	(0.974)
Intrinsic Motivation Math (School)	-15.62*	-3.085
	(8.044)	(5.579)
School Environment		
Disciplinary Climate	4.564***	1.538
	(1.112)	(0.949)
Disciplinary Climate (Peers)	22.45***	2.663
	(4.764)	(3.197)
Attitude Towards School	-1.425	-0.648
	(1.028)	(0.914)
Attitude Towards School (School)	9.319	6.164
	(6.337)	(4.444)
School Sense of Belonging	6.429***	7.356***
	(1.642)	(1.045)
School Sense of Belonging (School)	-1.445	14.18**
	(9.701)	(6.745)
Teaching Practices		
Math Self-Efficacy	20.52***	26.11***
	(1.313)	(1.069)

Math Self-Efficacy (School)	24.22***	10.04**
	(7.219)	(5.058)
Teaching Practices	9.415***	4.689***
	(1.149)	(1.138)
Teaching Practices (School)	19.16***	2.598
	(6.678)	(4.572)
Constant	175.4***	314.4***
	(48.46)	(32.81)
Observations	9,654	23,016
Constant	0.515	0.552

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1.

Table 1. A6. Math Scale: Detailed decomposition of Change in Scores in 2003-2012 by mean and percentiles

Variables	Mean	P 10	P 30	P 50	P 70	P 90
Individual						
Age	-0.0103	-0.0479	-0.0397	-0.0285	0.0369	0.0118
	(0.0199)	(0.0451)	(0.0288)	(0.0357)	(0.0379)	(0.0320)
Grade	-5.130***	-5.441***	-4.954***	-6.184***	-3.337***	-1.723***
	(1.111)	(1.200)	(1.075)	(1.343)	(0.730)	(0.384)
Girl	0.419*	0.283*	0.263*	0.388*	0.400*	0.455*
	(0.218)	(0.155)	(0.139)	(0.205)	(0.211)	(0.239)
Foreign Language	-0.343	-0.176	-0.175	-0.348	-0.491	-0.447
	(0.243)	(0.152)	(0.137)	(0.254)	(0.345)	(0.315)
Parental Highest Education (Years)	-0.165	0.0139	-0.0493	-0.171	-0.294	-0.188
	(0.120)	(0.0945)	(0.0702)	(0.140)	(0.209)	(0.145)
Immigrant First Generation	-0.765***	-0.517**	-0.670***	-0.957***	-0.881***	-0.579***
	(0.139)	(0.229)	(0.178)	(0.227)	(0.185)	(0.141)
Pre-School 2+ Years	0.255**	0.366**	0.146*	0.184*	0.184**	0.129*
	(0.120)	(0.179)	(0.0766)	(0.100)	(0.0907)	(0.0693)
Single Parent Family	-0.242***	-0.352***	-0.287***	-0.179*	-0.143*	0.0859
	(0.0741)	(0.121)	(0.0981)	(0.109)	(0.0849)	(0.0957)
Family Missing Information	-0.499***	-0.914***	-0.492***	-0.446**	-0.0919	0.0158
	(0.115)	(0.234)	(0.137)	(0.189)	(0.165)	(0.154)
Father Full Time Employed	-0.382***	-0.586**	-0.304**	-0.331	-0.484***	-0.370**
	(0.122)	(0.238)	(0.144)	(0.204)	(0.179)	(0.172)
Mother Full Time Employed	0.0841	0.0635	-0.0582	0.125	0.294**	0.0994
	(0.0572)	(0.0795)	(0.0603)	(0.105)	(0.118)	(0.0970)
Index of Economic, Social and Cultural Status	3.511***	0.823	1.341***	4.350***	5.079***	3.872***
	(0.608)	(0.506)	(0.394)	(0.834)	(0.904)	(0.739)
Peer Characteristics (School)						
Index of Economic, Social and Cultural Status	5.144***	2.948**	2.732***	7.506***	5.059***	3.971***
	(1.195)	(1.190)	(0.917)	(1.616)	(1.317)	(1.054)
Share of Immigrants First Generation	0.921*	-0.358	-0.118	1.506**	1.981***	1.338***
	(0.519)	(0.793)	(0.537)	(0.715)	(0.562)	(0.428)
School Characteristics						
Teacher Shortage	0.265	-0.0252	0.332	0.599	0.429	-0.0981
	(0.417)	(0.478)	(0.344)	(0.513)	(0.531)	(0.421)
Rural	-0.0686	-0.0975	-0.0653	-0.0723	-0.0755	0.0273
	(0.162)	(0.230)	(0.154)	(0.173)	(0.179)	(0.0710)
Public	-0.0347	-0.0597	-0.137	-0.0432	-0.0287	0.106
	(0.108)	(0.138)	(0.165)	(0.141)	(0.117)	(0.145)
Math Attitudes and Motivations						
Math Axiety	0.136	0.00907	0.0726	0.161	0.198	0.182
	(0.247)	(0.0210)	(0.132)	(0.293)	(0.359)	(0.331)
Math Anxiety (School)	-0.00530	-0.00153	-0.00298	-0.00283	-0.00771	-0.00952
	(0.209)	(0.0604)	(0.117)	(0.111)	(0.303)	(0.375)
Math Enjoying	-0.555***	0.118	0.0333	-0.322**	-0.947***	-1.412***
	(0.156)	(0.131)	(0.0914)	(0.155)	(0.267)	(0.367)
Math Enjoying (School)	0.410	-1.599**	0.0185	1.140	0.919	1.285**
	(0.535)	(0.785)	(0.496)	(0.759)	(0.614)	(0.613)
Intrinsic Motivation Math	-1.110***	-0.951***	-0.718***	-0.914***	-1.234***	-1.336***
	(0.157)	(0.282)	(0.170)	(0.255)	(0.237)	(0.227)
Intrinsic Motivation Math (School)	-2.337***	-1.521	-1.062	-3.811***	-1.890*	-1.480*
	(0.853)	(1.216)	(0.787)	(1.050)	(0.984)	(0.873)
School Environment						

Disciplinary Climate	0.168*	0.227	0.354**	0.143	0.201	-0.193
	(0.0941)	(0.151)	(0.144)	(0.149)	(0.125)	(0.123)
Disciplinary Climate (Peers)	0.284	0.337	-0.0410	0.482	-0.0803	0.763**
	(0.271)	(0.355)	(0.240)	(0.394)	(0.290)	(0.372)
Attitude Towards School	-0.0960	0.871***	0.451***	-0.427**	-0.842***	-0.551***
	(0.103)	(0.248)	(0.151)	(0.190)	(0.212)	(0.171)
Attitude Towards School (School)	0.789*	0.0144	0.457	1.064*	1.144**	0.507
	(0.457)	(0.514)	(0.384)	(0.615)	(0.528)	(0.468)
School Sense of Belonging	0.297**	-0.0157	0.0236	0.418**	0.610**	0.249*
	(0.142)	(0.0696)	(0.0452)	(0.202)	(0.287)	(0.129)
School Sense of Belonging (School)	-0.183	-0.314	-0.317	0.0905	-0.270	0.315
	(0.259)	(0.359)	(0.248)	(0.321)	(0.295)	(0.291)
Teaching Practices						
Math Self-Efficacy	6.448***	4.222***	4.141***	7.967***	6.590***	5.361***
	(0.747)	(0.681)	(0.542)	(0.951)	(0.803)	(0.677)
Math Self-Efficacy (School)	2.311**	0.151	1.285	1.282	3.936***	2.904***
	(0.911)	(1.217)	(0.836)	(1.096)	(1.166)	(0.921)
Teaching Practices	-1.500***	-1.046**	-1.363***	-2.121***	-1.835***	-0.312
	(0.296)	(0.515)	(0.337)	(0.474)	(0.455)	(0.368)
Teaching Practices (School)	-0.816	0.341	-1.071	-1.373	-0.475	-0.0390
	(1.073)	(1.433)	(0.953)	(1.347)	(1.295)	(1.072)
Year 2012 Estimated Score	489.4***	378.2***	430.0***	491.7***	548.9***	596.2***
	(2.087)	(1.851)	(1.581)	(2.593)	(2.100)	(1.641)
Year 2009 Estimate Score	488.4***	378.3***	432.8***	490.7***	546.5***	594.3***
	(3.678)	(5.997)	(4.668)	(4.973)	(4.013)	(2.078)
Difference	1.070	-0.0640	-2.776	1.027	2.353	1.930
	(4.229)	(6.277)	(4.929)	(5.609)	(4.529)	(2.648)
Explained Factors	7.200*	-3.235	-0.273	9.676**	13.65***	12.94***
	(3.678)	(3.666)	(2.796)	(4.632)	(3.778)	(2.974)
Unexplained Factors	-6.129*	3.171	-2.503	-8.649*	-11.30***	-11.01***
	(3.411)	(6.270)	(4.368)	(4.737)	(3.922)	(2.625)
Constant	138.9*	308.7	141.1	159.7	122.1	92.50
	(81.30)	(195.5)	(127.6)	(138.2)	(123.8)	(82.87)
Observations	32,670	32,670	32,670	32,670	32,670	32,670

Note: Robust standard error in parentheses and clustered at the school level. *** p<0.01, **p<0.05,*p<0.1.

References

AMERMUELLER, A (2004). PISA: What Makes the Difference? Explaining the Gap in Pisa Test Scores between Finland and Germany. *ZEW Center for European Economic Research Discussion Paper No. 04-004*.

ARUM, R., & VELEZ, M. (2012). Improving Learning Environments - school discipline and student achievement in comparative perspective. *Stanford: Stanford University Press*.

BARRERA-OSORIO, F., GARCIA-MORENO, V., PATRINOS, H.A, & PORTA, E. (2011). Using the Oaxaca-Blidner Decomposition Technique to Analyze Learning Outcomes Changes Over Time: An Application to Indonesia's Results in PISA Mathematics. *World Bank Working Paper 5584*.

CALERO, J. & J. O. ESCARDIBUL (2007): "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003". *Hacienda Pública Española, Vol. 183 (4/2007), pp. 33-66*.

DI NARDO, J., FORTIN, N., LEMEIUX, & T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica, Vol 64, No. 5, pp. 1001-1044*.

Fast and all. (2014), Does Math Self-efficacy Mediate the Effect of the Perceived Classroom Environment on Standardized Math Test Performance? *Journal of Educational Psychology (forthcoming)*.

FIGLIO, D. (2007). Boys named Sue: Disruptive children and their peers. *Education Finance and Policy, 2(4), 376–394*.

FIRPO, S., FORTIN, N., & LEMIEUX, .T (2009). Unconditional Quantile Regressions. *Econometrica, Vol. 7, No 3*.

FLOOK, REPETTI AND ULLMAN (2005), Classroom social experiences as predictors of academic performance. *Developmental Psychology, 04/2005; 41(2):319-27*.

GAIL H. and BETZ N. (1989), An Exploration of the Mathematics Self-Efficacy/Mathematics Performance Correspondence. *Journal for Research in Mathematics Education Vol. 20, No. 3 (May, 1989), pp. 261-273*.

GARCÍA PÉREZ. J.I., HIDALGO-HIDALGO, M., ROBLES-ZURITA J.A, (2011). Does grade retention affect achievement? Some evidence from PISA *Applied Economics, Vol. 46, Issue 2, pp. 1373-1392*.

GARCÍA PÉREZ. J.I., Hidalgo-Hidalgo, M., Robles-Zurita J.A, (2012). Diferencias regionales en rendimiento educativo en España: La familia lo explica todo? *UPO Working Paper*.

GUIO, J.M & CHOI, A (2013). Evolution of the School Failure Risk During the 2000 Decade in Spain: Analysis of PISA Results with a two-level logistic model. *Working Paper IEB 2013/17*.

HEMBREE, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21, 33-46*.

LOUNKAEW, K. (2013). Explaining urban-rural differences in educational achievement in Thailand: Evidence from PISA literacy datA *Economics of Education Review, Vol. 37, December 2013, pp. 213-225*.

MACHADO, J.AF. and J. MATA (2004), "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression". *Journal of Applied Econometrics, Volume 20, Issue 4, pages 445-466.*

MARKS, G. N. (2010). What aspects of schooling are important? School effects on tertiary entrance performance. *School Effectiveness and School Improvement, 21(3), 267–287.*

OECD (2010). PISA 2009 Results: Learning Trends. Changes in Student Performance since 2000.

OECD (2010). PISA 2009 results: What makes a school successful?

OECD (2012). PISA 2009 Technical Report.

OECD (2013). OECD Skills Outlook 2013: First Results from the Survey of Adult Skills.

OAXACA, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review, Volume 14, Issue 3.*

PATRINOS, H., MCDONALD, K., PARANDEKAR, S. (2012). Learning Outcomes in Thailand. What can we Learn from International assessments?. *The World Bank.*

REGE, M., TELLE, K., & VOTRUBA, M. (2010). Parental Job Loss and Children's School Performance. *Review of Economic Studies (2011), Vol 78, pp 1462-1489.*

RICHARDSON, F. C, & SUINN, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric datA *Journal of Counseling Psychology, 79,551-554.*

SILVER, E. A, AND M. K. STEIN. (1996). "The QUASAR Project: The Revolution of the Possible' in Mathematics Instructional Reform in Urban Middle Schools. *Urban Education 30, 4: 476 522.*

UK Department for Business Innovation and Skills (2012). *BIS Research Paper Number 81.*

VALENZUELA, J. (2013). Job Loss at Home: Children's Grades during the Great Recession in Spain. *WP.*

WATTS B. (2011), Relationships of mathematics anxiety, mathematics self-efficacy and mathematics performance of Adult Basic Education students. *ProQuest, UMI Dissertations Publishing.*

Word Bank (2013), Developing Skills for Innovative Growth in the Russian Federation, Chapter 3.

ZASLAVSKY, C. (2004). Fear of Math. *New Brunswick, New Jersey: Rutgers University Press, 1994.*

ZINOVYEVA, N., FELGUEROSO, F., VÁZQUEZ, P. (2013). Inmigration and Student Achievement in Spain: Evidence from PISA *SERIES, DOI 10. 10007.*

2. Overcoming the barriers: factors affecting achievement in schools and students from a disadvantaged environment

José Manuel Cordero
Francisco Pedraja
Rosa Simancas
University of Extremadura

Abstract

This paper is focused on studying *resilient* students, i.e., those who achieve good academic results despite the fact - they are negatively affected by an adverse socioeconomic environment. Specifically, our main aim is to identify the main determinant factors when explaining the results of Spanish students who participated in the PISA 2012 survey , both in maths and problem-solving using computers. Our results show that the possession of a computer at home is more relevant than the availability of computers in the school. Likewise, we have detected the existence of a positive and significant relationship between *resilience* and a reduced class size. These and other results can be useful for the design of educational policies in our country.

Key words

Education, PISA, Determinants of educational performance, Multilevel analysis, Educational policy.

INTRODUCTION

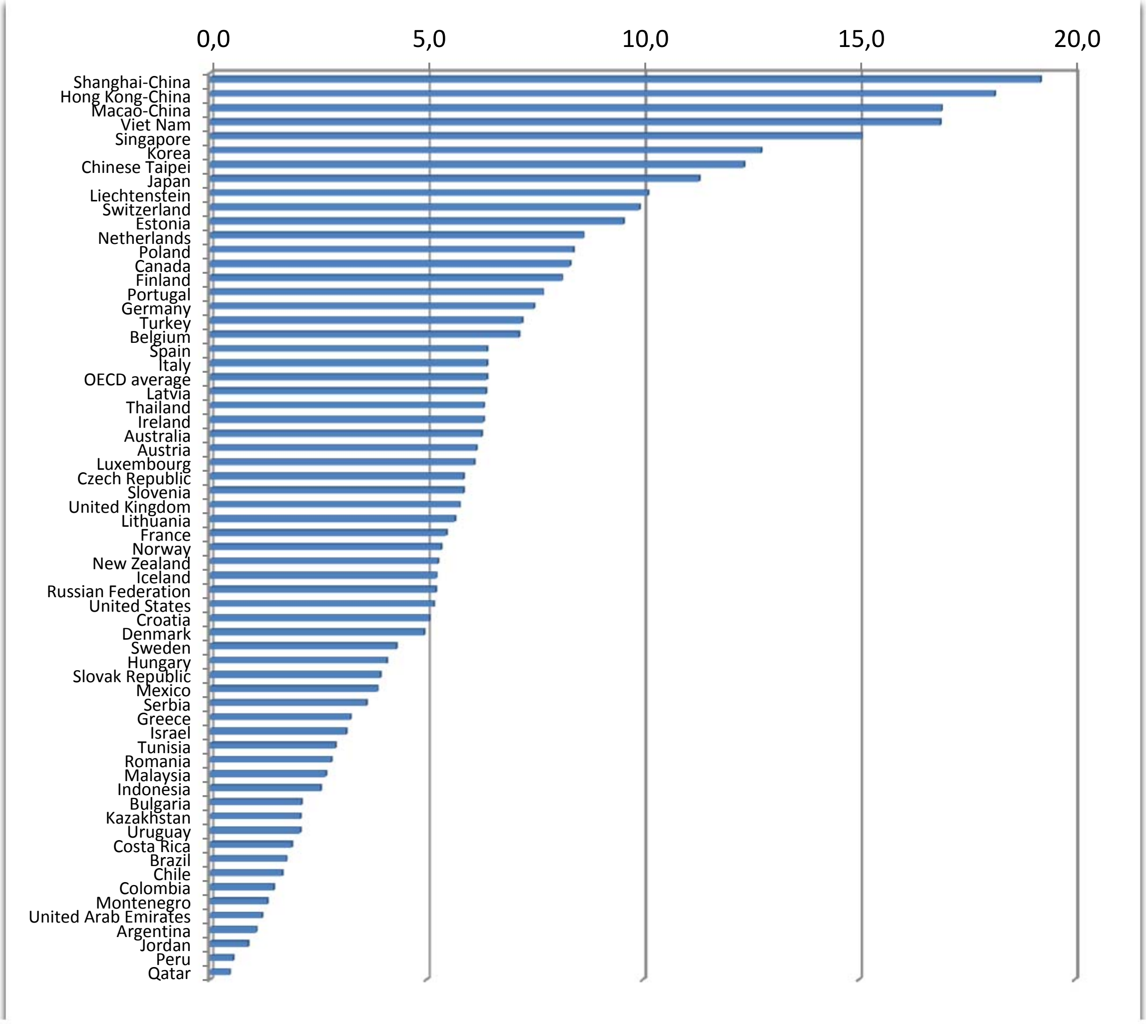
Since the first studies conducted in the field of Education Economics, one of the main concerns of researchers has been to investigate the determinants of academic performance (Colemanet al., 1966). In recent years, the answer to that question has been addressed from a comparative perspective thanks to the availability of an increasing number of international databases (Fuchs and Woessman, 2007; Hanushek and Woessman, 2011). A general conclusion of these studies is the relevant role of family socioeconomic background in explaining the academic performance of students (Sirin, 2005).

This factor is usually defined through indicators representing the educational level of parents, their level of job qualification and family wealth (Yang and Gustafsson, 2005). In the specific case of the Programme for International Student Assessment (PISA), developed by the Organization for Economic Cooperation and Development (OECD), this factor is approximated by the so-called index of economic, socialand cultural status (ESCS), composed of the highest educational and occupational levelof either of the parents and an indicator of cultural possessions at home. This variable shows a strong association with the academic performance of students. According to the data presented in the latest wave of the report (OECD, 2013a, p. 34), the differences in the ESCS index explain about 15% of the variation observed in maths scores among OECD countries.

The association between socioeconomic status of the student and educational results is so recognized that it is common to use its relationship with test scores as a measure of the degree of equity in education systems (Martins and Veiga, 2010; Rumberger, 2010). In this sense, an education system will be fairer and will better guarantee equal opportunities if it is effective in neutralizing the effects of the ESCS index of students on their school performance (Levin, 2010).

However, we should not assume that the existence of an association between these variables impliesthat the students from disadvantaged socioeconomic environments are inexorably condemned to school failure. Fortunately, there is a significant number of students who are able to overcome a disadvantaged socioeconomic background and achieve at high levels. These students, known in the literature as *resilients* (Wang et al., 1994), are the focus of our research. PISA 2012 identifies them as those who, though being in the bottom quarter of the ESCS in the country of assessment and performs in the top quarter across student from all countries after accounting for socioeconomic status. According to this definition, as shown in Figure 1.1, students fromAsian countries has the highest percentage of resilient students (between 15% and 20%), and Spain, with a percentage of 6.5%, presents a value very similar to the OECD average.

Figure 2.1. Percentage of resilient students of countries participating in PISA 2012



Source: OECD, PISA 2012 Database, Table II.2.7a

More specifically, our study attempts to identify the factors that characterize such students beyond their low socioeconomic status. For this purpose, we focus on the schools with students who belong to disadvantaged socioeconomic environments and, within them, we select students who achieve better academic results. Our aim is to find some common features between them, both related to their personal characteristics and the activities developed in their schools. With this approach we try to eliminate the effects related to the socioeconomic background of both the students themselves and the school, known in the literature as the'peer effect'[1], whose influence on performance is even higher than the student's own socioeconomic status (Willms, 2004). Ultimately, we intend to focus our analysis on those factors that characterize the resilient students, which can be modified through education policy measures with the goal of improving academic results.

Until now, previous research on this type of students has focused on identifying personal characteristics (Krovetz, 2007). Those studies generally agree in highlighting motivation or self-confidence as the main factors explaining the phenomenon of resilience (Wayman, 2002; Borman and Overman, 2004). However, we must not forget some school

1.Usually calculated from the average socioeconomic level of the classmates or schoolmates. For a review of peer effect studies see van Ewijk and Sleegers (2010).

factors that may also play a relevant role, as they are responsible for drawing our attention to other studies that insist on promoting regular attendance and participation in class (Masten and Coatsworth, 1998), maintaining a small number of students per classroom (Robinson, 1990) and per school (Noguera, 2002) or the application of innovative teaching practices that attempt to capture the attention of students from disadvantaged backgrounds and encourage them to develop their capabilities(Tajalli and Opheim, 2004).

In the United States there is a extensive literature devoted to the study of specific educational interventions for students at risk of school failure (Harris, 2007; Gregory et al, 2010), an aspect on which special emphasis has been placed since the adoption of the NCLB ("*No Child Left Behind*") Act in 2001. One of the main purpose of this regulation was to improve the results of the most disadvantaged students. Major advances in the characterization of resilient students have been made in the fields of psychology and sociology (Martin and Marsh, 2009), while contributions in the field of education economics are much scarcer. One exception is the study of Agasisti and Longobardi (2012) which, starting with the educational production function and through an econometric analysis, tries to identify some school characteristics with the existence of resilient students for the case of Italy using data from PISA 2009.

In our research we use a method similar to the previous study for the case of Spain using PISA 2012 data with respect to two competences: maths and problem solving. The first competence was the main core of the wave of 2012 (in 2009 it was the turn of reading competence and in 2015 it will be sciences). Almost two thirds of the assessment tests are devoted to this competence, as well as including a number of questions related to the specific attitude and disposition of students towards the subject. The test on problem-solving was performed by students using a computer (CBA, computer-based assessment), which leads us to wonder to what extent students who are more familiar with new technologies have a certain advantage over the rest for achieving better results[2]. Moreover, in the exploration of the factors that affect the performance of students in this competenceit it is possible to incorporate a set of variables reflecting non-cognitive aspects of attitude or perseverance in problem-solving drawn from an additional block included in the background questionnaire completed by students[3].

The procedure used in the selection of the analysis group is to segment the available sample to focus on schools with lower socioeconomic levels and, within them, to consider only those students with a socioeconomic level no higher than the bottom step that defines the segmentation of the selected schools. Thus, we attempt to isolate the socio-economic component of the analysis to focus our attention on other relevant factors both at an individual and a school level. Once the sample is segmented, we consider a student as resilient if the achieved score on each of the assessed competences is among the best ones in the distribution of results. Next, we estimate a multilevel logistic model that includes both individual and school variables as regressors in order to determine which variables are associated with the probability of belonging to the group of resilient students.

The rest of the study is organized as follows. The second section provides a description of the database used and a detailed explanation of the approach adopted to identify the schools and students under analysis. In the third section we explain the methodology used in the empirical analysis, i.e., the multilevel logistic regressions. The main results obtained from the estimations are presented and discussed in the fourth section. Finally, the article ends with the usual section of conclusions, in which we offer some recommendations for educational policy, based on the results obtained.

2. Marcenaro (2014) deals with this aspect in a greater detailin this volume.
3. For a specificanalysis of the non-cognitive aspects and theireffect on the problem-solvingcompetence seeMéndez (2014) in this volume.

DATABASE AND VARIABLES

The database used in our analysis comes from the PISA project (Programme for International Student Assessment), designed and launched by the OECD in the late nineties as a comparative, international, regular and continuous study of certain characteristics and competences of pupils aged 15 (Turner, 2006). Our research is based on the last wave, PISA 2012 and limited to the Spanish context, which provides information on a total of 25,313 students from 902 schools. The PISA 2012 report evaluates the performance of students in maths, reading comprehension, sciences and problem solving, going into greater depth in maths competence. This is the reason why in our study we shall use the results of this competence as a benchmark for identifying resilient students.

Likewise, in our empirical analysis we also analyze the results obtained in problem solving. The latter is defined by PISA 2012 (OECD, 2014, p.12) as "an individual's capacity to engage in cognitive processing to understand and resolve problem situations where a method of solution is not immediately obvious. It includes the willingness to engage such situations in order to achieve one's potential as a constructive and reflective citizen". Therefore, the aim of this independent evaluation is to assess whether the different educational systems are able to really prepare students to deal with the complex situations of daily life; due to that, the assessment of this competence does not measure the command of specific knowledge, but focuses on reasoning skills and the willingness with which the student confronts these problems that do not require prior acquired knowledge.

This knowledge was already assessed secondarily in PISA 2003 (OECD, 2005), from a focus closer to math problems. The biggest news from 2012 is that the tests have been carried out in digital format, which has facilitated the assessment and provided a greater volume of information. Now, the final score for each student not only reflects the result derived from their explicit response, but also includes information on the steps taken to reach that answer. Scores on problem solving are presented as a scale of skills represented by five plausible values, as in the competences of maths, reading comprehension and sciences (Wu and Adams, 2002)[4].

From the total sample of Spanish students participating in the PISA 2012 report, only a total of 10,175 students from 368 schools were evaluated in the problem solving competence by computer. Those students are the subject of this research since we can only make comparisons between the results in maths and problem solving for them. Table 2.1 shows the distribution of the sample of students and schools that participated in PISA problem solving compared to the total sample of the PISA 2012 survey by autonomic regions. As can be seen, there are two regions with a higher representation than the rest in the selected database, Catalonia and the Basque Country, especially the latter, whose students make up almost half of the sample. This is because these regions decided to participate with an enlarged representative sample in the specific computer-based assessment of competences which allowed them to conduct comparisons at an international level.

4. An extensive reviewof the plausible values can be found inMislevy (1991) or Mislevy et al. (1992).

Table 2.1. Spanish sample of students in PISA 2012 in problem solving by Autonomic Regions

	PISA 2012*		PISA problem solving	
	Students	Schools	Students	Schools
Balearic Islands	1,435	54	100	4
Cantabria	1,523	54	111	4
Castilla y León	1,592	55	201	7
Basque Country	4,739	174	4,739	174
La Rioja	1,532	54	85	4
Madrid	1,542	51	592	20
Galicia	1,542	56	202	8
Navarra	1,530	51	135	4
Murcia	1,374	52	141	6
Andalucía	1,434	52	910	33
Extremadura	1,536	53	150	5
Asturias	1,611	56	120	4
Aragón	1,393	51	159	6
Catalonia	1,435	51	1,435	51
Otros	1,095	38	1,095	38
Total	**25,313**	**902**	**10,175**	**368**

*Sample formed by students assessed in the basic competences: maths, reading comprehension and sciences.

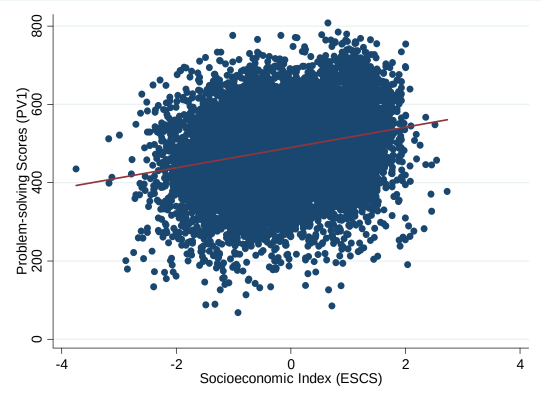
The results obtained by the Spanish students in both competences are similar, as shown in Table 2.2. The average score is below the OECD average in both cases, although the distance is greater in the case of problem solving, where the performance of Spanish students seems to be a little worse. If we look at the representative samples, the pattern is similar. In Catalonia the results are very similar, while in the case of the Basque Country they are somewhat better, getting above the OECD average in the case of maths.

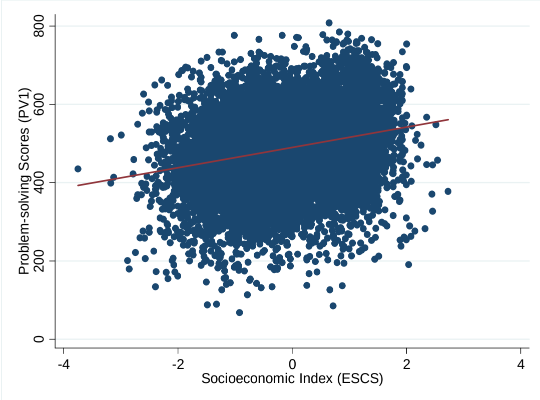
Table 2.2. Results (average scores) in maths and problem solving

		Maths		Problem solving	
Countries	**N**	**Average**	**Typ. dev.**	**Average**	**Typ. dev.**
S Korea	5,033	554	96.43	561	87.34
Japan	6,351	536	91.42	552	77.09
Canada	21,544	509	84.87	517	95.17
Australia	14,481	493	96.24	513	95.36
Finland	8,829	507	86.83	510	92.56
United Kingdom	4,185	498	92.72	520	92.44
Estonia	4,779	522	77.78	517	83.00
France	4,613	498	94.20	513	91.91
Netherlands	4,460	519	90.23	506	95.07
Italy	5,495	488	89.58	509	89.76
Czech Republic	5,327	520	94.42	528	90.14
Germany	5,001	514	94.27	508	97.92
United States	4,978	482	86.68	509	87.68
Belgium	8,597	520	99.07	514	101.52
Austria	4,755	508	88.14	508	88.52
Norway	4,686	489	87.03	502	96.67
Ireland	5,016	501	81.90	498	89.23
Denmark	7,481	486	83.62	481	92.78
Basque Country	4,739	508	80.41	498	92.60
Portugal	5,722	485	91.24	492	85.42
Sweden	4,736	479	88.26	492	89.73
Catalonia	1,435	497	81.08	491	97.15
Slovakia	4,678	486	99.15	487	95.03
Poland	4,607	521	88.42	483	93.24
Spain	10,175	496	83.55	488	96.87
Slovenia	5,911	485	87.42	458	96.90
Hungary	4,810	485	88.84	468	97.21
Turkey	4,848	449	90.06	456	75.23
Israel	5,055	469	99.70	456	117.12
Chile	6,856	445	84.47	466	85.45

Although the factors influencing these results maybe very different, usually the socioeconomic statusof families is identified as the most relevant.As mentioned in the introduction, this factor is approximated in PISA through the ESCS index, which takes the value 0 for the average of the OECD countries, so that negative values indicate a level below the average and positive values a higher level. Figure 2.1 shows the relationship between the results obtained in the two competences considered and the socioeconomic level of students. It is possible to identify a clear positive correlation in both cases, but with a greater dispersion in the case of problem solving.

Figure 2.1. Relationship between the socioeconomic level and the results in the two competences (total sample)





As anticipated in the introduction, the objective of this research is to isolate the effect of socioeconomic status in order to study the factors that characterize the students who obtain better results in unfavorable contexts. To achieve this goal, our strategy has been to segment the total sample in a way that from the 368 schools that participated in PISA problem solving; only those with a lower average socioeconomic status have been selected. In this first selection, we are left with the bottom third (33rd percentile) in terms of the ESCS variable[5]. Specifically, we have 125 schools with a total of 3,116 students available for our analysis. Then, to ensure that our study only includes students from an adverse socioeconomic environment, we select only those students whose individual socioeconomic level does not exceed the criterion used in the selection of schools, reducing the sample to 2,054 observations. Finally, we have decided to discard those schools with a small number of students (less than 10), so that the final sample used in our empirical analysis consists of 1,917 students from 105 schools.

As expected, the results obtained by the students belonging to this segmented sample are very low. Specifically, the average is 453 points in the case of maths and 449 in problem solving. These lower average values can be explained largely by a series of variables linked to socioeconomic level of the selected students. So, for example, in Tables 2.3 and 2.4 we can see the differences between the total sample (10,175 pupils) and the segmented sample (1,917 pupils) regarding the educational level of their parents and the number of books at home. The first one is characterized by having an average socioeconomic status, in which both a majority of fathers and mothers are university educated and half of households have more than 100 books, while the second has a very negative average value of the ESCS index (-1.12), parents, on average, do not have studies beyond compulsory secondary education (only 10% have university studies) and they have a small number of books at home (less than 100 in 80% of cases).

5. Although the PISA definition given in the introduction considers that resilient students are those found in the bottom quartile in terms of the ESCS varaible, we have preferred to opt for the bottom third in order to reduce the loss of observations. This same criterion is the one followed by Agasisti y Lomgobardi (2012).

Table 2.3. Educational levels of the parents of the students assessed in PISA 2012 for the total sample and the segmented sample selected in the study

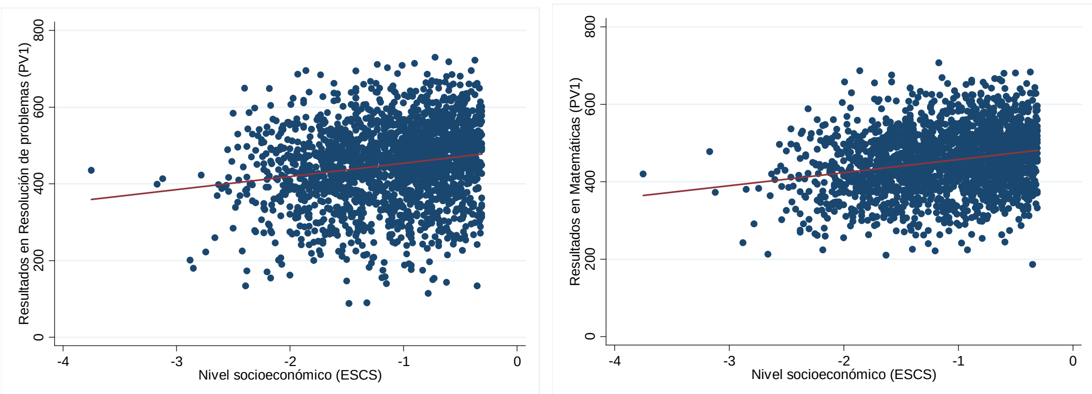
	Total sample		Segmented sample	
	Educational level of the father	Educational level of the mother	Educational level of the father	Educational level of the mother
None	2.44 %	1.80 %	7.07 %	5.94 %
Primary education	8.70 %	6.96 %	22.16 %	19.68 %
Secondary education	21.78 %	20.33 %	37.59 %	40.46 %
Baccalaureate or FP I	21.87 %	25.89 %	21.54 %	24.59 %
University studies	45.22 %	45.02 %	11.65 %	9.32 %

Table 2.4. Distribution of the number of books in the home in percentages, for the total sample and the segmented sample selected in the study

Books at home	Total sample	Segmented sample
0-10	6.99 %	18.35 %
11-25	12.30 %	23.47 %
26-100	29.85 %	35.86 %
101-200	22.08 %	13.92 %
201-500	17.66 %	6.22 %
More than 500	11.13 %	2.16 %

With a much more homogeneous subsample regarding the socioeconomic level of the families, the relationship between the results in the two competences and the ESCS index becomes much weaker, as can be seen in Figure 2.2., which allows us to concentrate on the study of other variables related to the results.

Figure 2.3. Relationship between the socioeconomic level and the results after the segmentation



Specifically, as previously stated, our goal is to identify the resilient students within the sub-sample of the most disadvantaged schools and students in terms of the ESCS variable, i.e., those who achieve better scores in either of the two competences considered. To do this, we generate two dummy variables that will become the dependent variables of our models. The first is *Resilientin Maths*, which takes the value1 if the student is in the top quartile of the distribution of results in maths and the second dependent variable is *Resilient in Problem Solving*, coded analogously to the above[6]. Tables 2.5 and 2.6 present some descriptive statistics in terms of socioeconomic level and results that allow us to characterize both groups.

Table 2.5. Descriptive statistics referring to maths competence

Variables	*Resilient* in maths				Not *resilient*			
	Average	Typ.dev.	Min	Max	Average	Typ.dev.	Min	Max
PV1MATHS	557.87	40.75	508.19	707.21	418.20	59.68	186.41	508.11
PV2MATHS	551.81	49.30	441.04	726.99	419.73	63.49	175.50	570.11
PV3MATHS	551.03	48.23	407.63	726.99	419.65	64.27	154.47	571.90
PV4MATHS	552.64	49.15	438.08	730.81	420.49	63.31	184.07	561.08
PV5MATHS	551.27	48.18	411.44	743.35	419.83	62.93	154.47	570.42
ESCS	-0.97	0.48	-2.46	-0.31	-1.18	0.55	-3.75	-0.31

Table 2.6. Descriptive statistics referring to problem solving competence

Variables	*Resilient* in problem solving				Not *resilient*			
	Average	Typ.dev.	Min	Max	Average	Typ.dev.	Min	Max
PV1CPRO	543.91	72.00	275.74	729.44	416.55	93.36	87.44	662.08
PV2CPRO	539.07	75.55	273.30	743.24	418.10	96.01	73.64	689.67
PV3CPRO	538.19	74.97	265.19	776.52	417.92	96.37	67.96	687.24
PV4CPRO	539.25	74.52	279.80	777.33	419.02	95.29	66.34	690.48
PV5CPRO	539.29	75.66	284.67	748.11	417.49	95.42	59.03	692.11
ESCS	-0.97	0.48	-2.46	-0.31	-1.18	0.55	-3.75	-0.31

These data allow us to identify huge differences in educational results between the groups of individuals in the two competences considered, tough they have a similar socioeconomic level on average. This encourages us to investigate the potential existence of other possible factors that explain such discrepancies in results, both with regard to individual characteristics of students and those aspects related to the school. In both cases we shall distinguish between general and specific explanatory variables of each competence, the latter referring to variables linked solely to one of the two competences assessed and, therefore, only included in the estimation of the corresponding model.

Firstly we have selected three control variables that, although not directly related to socioeconomic aspects, we could expect to have some influence on the dependent variable according to the previous literature on the determinants of students'performance. These are gender, represented by a dummy variable that takes the value 1 if the student is a girl, first-generation immigrant status and family structure, also represented by a dummy variable that takes the value 1 if the student is part of what is known as the traditional family, i.e.,that formed by both parents and their children.

Besides these control variables and considering the main goal of this study, we have tested the possible incorporation into the model of a number of individual indicators related to educational quality of the school. In particular, we have included an index

6. Both dependent variables have been generated based on the first plausible value of each competence.

representing the disciplinary climate in class, constructed from the student responses about the frequency with which disruptions occur. Also, our interest in testing the influence of computer resources has led us to incorporate the possession of computer at home as a possible explanatory variable.

The next block of variables is made up of those which in principle should be related to scores in maths. Among them, we selected several dummies such as the student's enjoyment with maths, attention paid in class and the effort shown by friends in the classes on this subject. Meanwhile, among those linked to the problem solving competence, we consider whether the student has a computer available at the school and also claims to use it for educational purposes together with two continuous indices developed by PISA experts according to student´ responses to different questions related to non-cognitive aspects. The first includes the disposition or attitude with which the student faces a problem (*Openness to problem solving*), while the second tries to approximate the determination and perseverance of the student in that situation (*Perseverance*). In both cases, a higher value of the indices should be correlated with greater skills in problem solving.

Among the school variables we can also distinguish some that are general and others specifically linked to each of the assessed competences. Within the first block, several indicators are included, obtained from the responses of principals on the degree of autonomy with which the school operates (responsibility of the school for hiring and firing teachers, setting their salaries and wage rises, or formulating and allocatingthe school budgets) or the quality of educational resources (availability of computers for educational purposes, educational software, calculators, books, audio-visual resources and laboratory equipment)[7].

Two variables related to teachers have also been incorporated: the teacher-student ratio and the relationship between the teaching staff and students of the school. Likewise, the level of absenteeism registered in the school, obtained from the opinion of the principal about the regularity with which students attend classes, and the average class size in the school were also considered. In this latter case, after analyzing the frequency distribution among the schools which make up the analyzed sample, it has been set to a value of less than 20 students in order to select schools with small class size.

Finally, the selection of specific variables has been limited to one per competence. Thus, hours of math instruction, obtained using a continuous variable reflecting the average weekly time (in minutes) of maths classes, is included as an indicator that should be associated with obtaining better results in this subject; on the other hand, the number of computers available for instruction has been selected as a variable that could have related to the results obtained on the computer-based test in problem solving.

Table 2.7 shows the main descriptive statistics of all the variables considered in our analysis, distinguishing between dependent, individual and school variables.

7 In order to facilitate the interpretation of the parameters associated with these indices, they have been transformed into dummy variables that take thevalue 1 if the schools are at the very top of the distribution in each case.

Table 2.7. Descriptive statistics of the variables included in the empirical analysis

VARIABLES **Dependent Variable**	**Minimum**	**Maximum**	**Average**	**Typ. dev.**
Resilient in maths	0.00	1.00	0.2499	0.4331
Resilient in problem solving	0.00	1.00	0.2587	0.4381
Student Level Variables				
General				
Gender	0.00	1.00	0.5013	0.5001
Immigrant	0.00	1.00	0.1607	0.3673
Family	0.00	1.00	0.8164	0.3873
Computer	0.00	1.00	0.9259	0.2620
Climate	-2.48	1.85	-0.1196	0.8607
Specific				
Enjoys Maths	0.00	1.00	0.2374	0.4256
Peer Maths	0.00	1.00	0.2796	0.4489
Attention Maths	0.00	1.00	0.5467	0.4979
Attitude	-3.63	2.45	-0.0963	0.7904
Perseverance	-4.05	3.53	0.0290	0.7471
School PC	0.00	1.00	0.6427	0.4793
School Level Variables				
General				
Autonomy	0.00	1.00	0.2358	0.4246
Teacher-student Ratio	0.00	1.00	0.3292	0.4700
School resources	0.00	1.00	0.2932	0.4553
Absenteeism	0.00	1.00	0.3933	0.4886
Teacher-student Relationship	0.00	1.00	0.1137	0.3176
Small class	0.00	1.00	0.2306	0.4213
Specific				
Hours Maths	157.14	298.08	210.2599	27.1170
Nº Computers	12.00	200.00	43.0498	26.9030

The values of the descriptive statistics allow us to detect that there are hardly any differences in the gender composition of the sample. The percentage of immigrant students in the sample (16.1%) is significantly higher than the 9.9% registered in the national sample for PISA 2012 (INEE, 2013), a result that could be expected given the link existing between the immigrant condition and low socioeconomic status. Moreappealing is the low proportion of students who claim to enjoy maths or have friends who put effort and determination in to this subject and the high level of absenteeism, considering that the variable refers to schools where students do not attend class regularly and not sporadically. As for the rest of the school variables, perhaps the most striking result is that 23% of schools have classes with quite a small average size (less than 20 students).

METHODOLOGY

The model used inthe empirical applicationis a multilevel regression (Bryk and Raudenbush, 1992; Goldstein, 1995), which considers that the students are grouped (nested) at a higher level represented by the schools. With this technique possible biases in the estimations derived from the correlation between the values of the school variables of pupils from the same school (Hox, 2002) are avoided. Since the dependent variables are categorical, these regressions assume a binomial logistic model structure.

This approach has previously been used in different studies using the PISA database to analyze the main factors related to the probability that a particular situation may occur, such as school failure (Calero et al, 2010; Cordero et al, 2012) or grade retention (Goos et al, 2012; Carabaña, 2013; Cordero et al, 2014).

In this model, the dependent variable represents the group of students with higher scores in PISA (first quartile within the selectedsub-sample), where the variable to be estimated is the probability that the student "*i*" from school "*j*" is included within the corresponding group: $P(Y_{ij} = 1 | \beta) = P_{ij}$. This probability can be modeled using the following logistic function:

$$\log\left[\frac{P_{ij}}{(1-P_{ij})}\right] = \beta_{0j} + \beta_{ij} X_{ij} + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_1 \qquad (1)$$

In this equation, the probability that the student meets the established requirement depends on a vector of independent variables at the individual level (Xij) and a vector of school variables (Zj), but also takes into account the deviation of the school *j* (uj) with respect to the results performance of all the schools (γ0) and the deviation of student *i*with respect to the average of the results obtained by students who belong to the same school *j*.

The values of the estimated coefficients in the model can not be interpreted directly as in a linear regression, so we need to estimate the odds ratios of each independent variable. These statistics measure the relationship between the probability of an event happening over the probability if it not happening when the value of the variable considered increases by a unit, keeping the other ones constant. Therefore, the odds ratios associated with an explanatory variable will take a value greater than one if that variable increases the probability that a student belongs to the group of those with higher academic performance, and less than one if that variable decreases the probability of such an event. Thus, the former will be associated with positive coefficients and the latter with negative coefficients.

The most commonly used strategy for the calculation of the results in this type of study is the use of an "additive" approach where, from a basic starting specification, they consider the different blocks of explanatory variables step by step (Dronkers and Robert,

2008) incorporating, firstly, the variables related to student level, and subsequently, the variables corresponding to the school level.

ANALYSIS AND DISCUSSION OF RESULTS

In this section we present the results obtained by applying the multilevel logistic regression model to the sample of students selected according to the criteria explained above. In order to carry out this estimation, the problem of the lack of responses in some of the variables (missing data) has been addressed using the method of imputation by regression recommended by the OECD (2008). The estimations were performed using the HLM 6 software (Raudenbush et al., 2004), with which it is possible to incorporate the sample weightings in the estimations to ensure that sampled students adequately represent the analyzed total population (Rutkowski et al., 2010)[8]. Thus, the results of the analysis refer to the entire Spanish population, despite there are some regions (Basque Country and Catalonia) with a greater representation in the sample because of having participated with an enlarged sample.

In the estimation we have considered two alternative models depending on the competence assessed. In the first model, the dependent variable takes the value 1 if the student is considered as a resilient in the maths competence, while in the second model, the same dependent variable represents resilience according to the student' results in problem solving.

The estimation of these models has followed a sequential structure, just as described in the above section. Thus, firstly, only variables at an individual level have been incorporated into the analysis, distinguishing between general and specific variables for each competence. This distinction is necessary because there are variables associated with each of these competences, such as the level of enjoyment in maths or the degree of identification with the levels of perseverance and preference for complex problem solving. Logically, each of these specific blocks has only been incorporated into the estimation of the model corresponding to the dependent variable that they are linked to. The results of these estimations are shownin Table 2.8.

8 These weights include adjustments for non-response by some schools and students within schools to prevent a small set of schools or students can be overrepresented. These processes are based on intensive calculation methods, known as resampling methods, which consist of taking multiple samples from the original sample. Specifically, PISA uses the Balanced Repeated Replication (BRR) with 80 replicates. For an extensive description of this procedure, see (OECD 2009)..

Table 2.8. Result of the estimations with variables at the student level

	Maths					Problem Solving			
VARIABLES	**Coeff.**	**SE**	**Odds Ratio**		**VARIABLES**	**Coeff.**	**SE**	**Odds Ratio**	
Constant	-1.98	0.52	0.14	***	Constant	-2.51	0.38	0.08	***
Student Level					**StudentLevel**				
General					**General**				
Gender - female	-0.49	0.13	0.61	***	Gender - female	-0.34	0.13	0.71	***
Immigrant	-0.86	0.30	0.42	***	Immigrant	-0.86	0.26	0.42	***
Family	0.59	0.17	1.80	***	Nuclear family	0.65	0.19	1.92	***
Computer	0.81	0.34	2.24	**	Computer	0.88	0.33	2.42	***
Climate	0.28	010	1.33	***	Climate	0.30	0.10	1.35	***
Specific					**Specific**				
Enjoys Maths	0.29	0.16	1.25	*	Openness	0.62	0.09	1.85	***
Peer maths	0.23	0.16	1.26		Perseverance	0.15	0.08	1.16	**
Attention maths	-0.03	0.14	0.97		SchoolPC	0.30	0.16	1.36	**

*** indicates that the variable is significant to 99%; ** to 95%; * to 90%.

The main conclusion from this first analysis is that no significant differences are detected between the parameters associated with the explanatory variables in the two alternative models considered, so we can interpret both of them together.

As expected, the three control variables incorporated into the analysis have a significant influence and in the same direction on the dependent variable in both competences. Both female gender and immigrant status are negatively linked to the probability that the student belongs to the group of resilient students, while belonging to a traditional family has the opposite effect. These results are in line with previous literature, since there are many studies that have identified these factors as good predictors of results. However, we must insist on the fact that what our dependent variable is the possibility of coming among the top students in a context characterized by students from disadvantaged socioeconomic backgrounds.

One result which we would like to draw attention to is the high, as well as clearly significant, value registered in the odds ratios representating the possession of a computer at home in both models. According to these parameters, this factor seems to be the most relevant of all the individual factorsin explaining resilience. This result contrasts with that obtained by Mediavilla and Escardíbul (2014) in this volume, who do not find a significant relationship between the possession of computer at home and results. However it is worth noting that our analysis is referred to a sub-sample, those with worse scores in the ESCS index, while their study refers to the entire sample of Spanis students who participated in the test. One possible explanation for this discrepancy could be that the influence of having computer at home disappears when taking into account other variables representing socioeconomic status, while if these factors are discounted in the assessment, it becomes apparent that the presence of a computer at home is a relevant factor in explaining academic success.

On the other hand, the perception of students regarding discipline in the classroom (climate) is also positively and significantly associated with the likelihood of achieving academic success, although its relative importance is much lower. This result agrees with that obtained in the study of Padron et al. (1999) which is devoted to resilient students in primary education and concludes that these students are more aware of an atmosphere of learning and spend more time interacting with teachers on aspects related to teaching than other students.

Turning to the results obtained relating to specific variables of attitudes shown towards maths and problem solving, we find some differences between the two competences. In the case of problem solving, there are more variables with significant values while only one variable is significant, with a low level of confidence, in maths. According to the value of its odds ratio, the most relevant variable, positively associated with resilience in problem solving, is represented by student attitudes when they confront this type of problems. However, this result should be interpreted with caution since this configuration of the variable may reveal a problem of endogeneity with the dependent variable due to the existence of a relationship of inverse causality, because, although students with better problem-solving attitudes are those that get better scores, it is also true that better results can be achieved by those who have a more positive attitude when tackling a problem.

The availability of computers at school and theiruse for educational purposes also play a prominent role. This result is similar to that obtained by Cabras and Tena (2013) using information on the maths test in PISA 2012. They conclude that there is a positive causal effect between computer use and students' school performance, especially of those coming from a more disadvantaged socioeconomic background.

After exploring the associations between different indicators selected individually and the two dependent variables, in the next step of the empirical analysis we add the variables related to the school. Again, in the study we incorporate a set of both common and specific variables for each competence. In the latter case we refer to the number of hours of maths taught each week which is only included in the first model, while the number of computers available is exclusively considered in the model regarding the problem solving (as mentioned, such problems are resolved using the CBA method). The results obtained in this new estimation are presented in Table 2.9.

Table 2.9. Result of the estimations with variables at the student and school level

Maths					Problem Solving				
VARIABLES	**Coeff.**	**SE**	**Odds Ratio**		**VARIABLES**	**Coeff.**	**SE**	**Odds Ratio**	
Constant	-0.87	0.65	0.42		Constant	-3.09	0.41	0,05	***
Student Level					**StudentLevel**				
General					**General**				
Gender - female	-0.51	0.13	0.60	***	Gender - female	-0.34	0.13	0,71	***
Immigrant	-0.93	0.31	0.39	***	Immigrant	-0.88	0.27	0,41	***
Family	0.59	0.18	1.80	***	Nuclear family	0.63	0.20	1,87	***
Computer	0.84	0.35	2.32	**	Computer	0.93	0.33	2,53	***
Climate	0.30	0.11	1.36	***	Climate	0.30	0.11	1,35	***
Specific					**Specific**				
Enjoys Maths	0.32	0.17	1.28	**	Openness	0.65	0.09	1,91	***
Peer maths	0.23	0.17	1.26		Perseverance	0.16	0.09	1,17	**
Attention maths	-0.04	0.15	0.96		SchoolPC	0.29	0.17	1,34	**
School level					**Schoollevel**				
General					**General**				
Autonomy	0.24	0.17	1.27		Autonomy	0.44	0.25	1,55	*
Teacher-student Ratio	0.43	0.20	1.54	**	Teacher-studentRatio	0.19	0.20	1,21	
School resources	0.07	0.21	1.07		Schoolresources	0.07	0.21	1,07	
Absenteeism	-0.34	0.17	0.71	**	Absenteeism	-0.35	0.20	0,70	*
Teacher relationship	1.03	0.30	2.80		Teacherrelationship	0.93	0.27	2,54	
Small class size	0.97	0.24	2.63	***	Small classsize	1.08	0.27	2,94	***
Specific					**Specific**				
Hours Maths	-0.01	0.00	0.99	***	Nº computers	0.00	0.00	1,00	

*** Indicates that the variable is significant to 99%; ** to 95%; * to 90%.

In general, most of the parameters associated with the individual variables are maintained when incorporating the school variables, and for this reason the discussion about the results of this new model will focus on these latter variables. A first consideration to highlight is that in this new estimation we do see differences between the parameters associated with resilience in the two competences assessed, although there are common patterns that lead us to a series of very interesting conclusions on indicators relating to the school context, which we are going to place further emphasis upon the next lines.

The factor showing a higher level of correlation with the two dependent variables is that of belonging to a small class, with a weight which is significantly greater than the other variables[9]. Although the analysis carried out does not allow us establish causal relationships between class size and academic success[10], this result is consistent with evidence found in other studies, which indicate the major influence of class size in schools with more disadvantaged socioeconomic environments (Heinesen, 2010). In these cases, the possibility of having fewer students per class favors more personalized attention that,

9.The possibility of being able to distinguish between class size and the teacher-student ratio is a great advantage compared to previous waves of the Report, in which only this latter indicator was available. This indicador usually reflects the number of students with special needs in the school more than the class size (Hanushek, 1999).
10.See Chingos (2013) for a recentreview of the literatureonthis issue.

to some extent, can compensate for the fact that their parents can not give them as much help at home as in more favorable contexts (Fredriksson et al., 2014).

Another factor that has a significant influence on both models is school absenteeism, though logically, in this case, the effect has the opposite direction, a result in line with the empirical evidence existing on an international scale (OECD, 2013b). This phenomenon has led to the proposal of different strategies to encourage class attendance, but with still relatively little success (Reid, 2013).

If we look at the two models separately we can also see some interesting results. For example, in the case of the problem-solving competence, it can be seen that the number of computers available for instruction has no significant effect, which coincides with the result of other previous studies (Calero and Escardíbul, 2007;Cordero et al, 2012) in which they use information on the Spanish students in previous waves of PISA. However, the individual variable representing the use of computers in the school for educational purposes does show a positive and significant correlation with the probability of achieving good results in the problem solving competence, as has been shown previously. In view of these results, it could be argued that policies based on the indiscriminate increase in the provision of computers in schools have no effect on the academic results of students, unlike what could be expected when these improvements in provisions are accompanied by a strategy to promote their use for educational purposes.

The degree of autonomy with which the schools operate seems to reflect a correlation, albeit weak, with the academic results of students in problem solving. This relationship, however, does not appear to be statistically significant in the case of maths, which makes us be cautious about this result. In fact, in previous studies which have analyzed this question with the same PISA database, contradictory results have been found. While Maslowski et al. (2007) did not find an association between levels of autonomy and academic performance in an international analysis, Hanushek et al. (2013) does detect a relationship between the two, though only in the most developed countries. Although, it is common to detect significant differences in the degree of autonomy in international comparisons, we should not be surprised by our result in Spain where there is very little room for the schools to put this autonomy into practice.

Although in principle the lack of significance of the variable representing the relationships between students and teachers in both specifications of the model may surprise us, the fact is that this result is largely explained by taking into account that it is based on the opinions of principals regarding these relationships, which might entail a significant bias when correctly identifying the effect that we are trying to identify.

Finally, the hours of maths instruction is clearly significant although its odds ratio, with values close to 1%, alerts us to an insignifican impact on results. This reflects a difference between the Spanish case and the international evidence, in which this factor is often associated with better academic performance of students, especially those belonging to the more developed countries (Lavy, 2012).

CONCLUSIONS

In this study we have carried out an analysis of the determinants of school performance of Spanish students considered as resilient in two of the competences assessed in PISA 2012, maths and problem solving by computer. These students are those who, despite coming from a socioeconomically adverse background, are able to succeed in term of academic achievement.

Knowing the importance of the socioeconomic environment in explaining academic results, we have sought to isolate this component by selecting only those schools that have the lowest average socioeconomic level for the analysis, so that the effect of other factors, that are normally limited by the influence of that environment, could be shown more clearly.

Despite the fact that the students from these schools have, a priori, many chances of obtaining low scores in PISA, it is possible to identify some factors that characterizes more successful schools. According to our results, these schools provide teaching in classrooms with smaller class sizes (less than 20 students), they maintain a certain level of discipline (good climate), with low rates of absenteeism and regular use of computers for instruction. All these variables seem to be clearly related to the quality of teaching, which confirms this factor as a key element in motivating students to be able to overcome the harsh conditions of a disadvantaged socio-economic background and to be able to develop their potential (Hanushek, 2011).

These results provide us with some keys to the design of educational policies aimed at those schools with students coming from relatively low socioeconomic levels. In this sense, it would be useful to increase the teaching staff so that there would be fewer students per classroom or, alternatively, establish some kind of system of incentives (positive or negative) for both teachers and principals of schools to encourage regular student class attendance as well as a regular use of computers in the teaching of some subjects.

Among the individual variables considered, immigrant status and the fact of being a girl are those that have a more significant influence (negatively) on the success achieved by students from more disadvantaged backgrounds. In the latter case, as is already known from other studies, this result can be explained by the type of competence assessed (maths)[11]. Anotherinteresting result is the high correlation observed between the presence of a computer at home and resilience, an association which disappears when the analysis includes all students, evidence that has been proved in other studies (Marcenaro, 2014).

In any case, the results obtained and discussed above should be used with caution due to the fact that they are based on an analysis performed usingcross-section data, thus they can not be interpreted in a causal sense. In this regard, we believe that the decision-making for educational policy must be based on the evaluation of policies using causal inference techniques that allow us to measure more precisely the possible impact of the implemented policies (Angrist and Pischke, 2008; Schlotter et al., 2011). In this sense, it would be ideal to conduct random or controlled trials, however the high cost of these initiatives leads us to recommend that, at least, efforts should be intensified to develop

11. For a more detailed analysis of gender differences see Mediavilla and Escardíbul (2014) in this volume.

longitudinal databases that allow us to evaluate the impact of particular educational interventions over time, such as the modification of class sizes (Fredriksson et al., 2013).

REFERENCES

AGASISTI, T. and LONGOBARDI, S. (2012)"Inequality in education: can Italian disadvantage students close the gap? A focus on resilience in the Italian school system".Documento de Trabajo del IEB, 2012/39.

ANGRIST, J. D., PISCHKE, J. S. (2008) Mostly harmless econometrics: An empiricist's companion, Princeton University Press.

BORMAN, G. D. and OVERMAN, L. T. (2004) "Academic resilience in mathematics among poor and minority students". The Elementary School Journal, 177-195.

BRYK, A.S. and RAUDENBUSH, S.W. (1992) "Hierarchical Linear Models: Applications and Data Analysis Methods".SagePublications, Newbury Park, ThousandOaks, CA.

CABRAS, S. and TENA, J.D. (2013)"Estimación del efecto causal del uso de ordenadores en los resultados de los estudiantes en la prueba PISA 2012".En INEE (ed.): PISA 2012: Programa para la evaluación de los alumnos. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

CALERO, J., CHOI, A. and WAISGRAIS, S. (2010)"¿Qué determina el fracaso escolar en España? Un estudio a través de PISA 2006". Revista de Educación, nº extra 2010: 225-256.

CALERO, J. and ESCARDÍBUL, J. O. (2007)"Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003". Hacienda Pública Española, 183, 4, 33-66.

CALERO, J. and ESCARDIBUL, O. (2013)"El rendimiento del alumnado de origen inmigrante en PISA 2012".En INEE (ed.): PISA 2012: Programa para la evaluación de los alumnos. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

CARABAÑA, J. (2013)"Repetición de curso y puntuaciones PISA ¿Cuál causa cual?". En INEE (ed.): PISA 2012: Programa para la evaluación de los alumnos. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

CHINGOS, M. M. (2013). "Class size and student outcomes: Research and policy implications". Journal of Policy Analysis and Management, 32(2), 411-438.

COLEMAN, J. Et al (1966) "Equality of Educational Opportunity".Washington: Office of Education.

CORDERO, J.M. Et al (2012)"Análisis de los condicionantes del rendimiento educativo de los alumnos españoles en PISA 2009 mediante técnicas multinivel". Presupuesto y Gasto Público, 67, 71-95.

CORDERO, J.M. Et al(2014)"La repetición de curso y sus factores determinantes en España".Revista de Educación, en prensa.

DRONKERS, J. and ROBERT, P. (2008) "Differences in Scholastic Achievement of Public, Private Government-Dependent and Private Independent Schools". Educational Policy, 22 (4), 541-577.

FREDRIKSSON, P.; ÖCKERT, B. and OOSTERBEEK, H. (2013) "Long-term effects of class size". The Quarterly Journal of Economics, 128 (1), 249-285.

FREDRIKSSON, P.; Et al (2014) Inside the Black Box of Class Size: Mechanisms, Behavioral Responses, and Social Background. Institute for the Study of Labor (IZA), DP No. 8019.

FUCHS, T. and WOESSMANN, L. (2007) "What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data". Empirical Economics, 32 (2-3): 433-464.

GOLDSTEIN, H. (1995)"Multilevel statistical models (2nd edition) ". New York: Wiley Publishers.

GOOS, M. Et al (2012)"How Can Cross-Country Differences in the Practice of Grade Retention Be Explained? A Closer Look at National Educational Policy Factors". Comparative Education Review, 57 (1), 54-84.

GREGORY, A. Et al (2010)"The Achievement Gap and the Discipline Gap: Two Sides of the Same Coin?". Educational Researcher, 39(1), 59-68.

HANUSHEK, E.A. (1999) "The evidence on class size. In Mayer, S.E., Peterson, P.E. (eds.). Earning and learning: How schools matter". Washington, DC: Brookings Institution Press.

HANUSHEK, E.A. (2011)"The economic value of higher teacher quality". Economics of Education Review, 30(3), 466-479.

HANUSHEK, E.A. and WOESSMAN, L. (2011) "The economics of international differences in educational achievement".En Hanushek, E.A., Machin, S., Woessmann, L. (eds). Handbook of the Economics of Education, vol. 3, Amsterdam: North Holland, pp. 89-200.

HANUSHEK, E. A. Et al (2013)"Does school autonomy make sense everywhere? Panel estimates from PISA". Journal of Development Economics, 104, 212-232.

HARRIS, D.N. (2007) "High-flying schools, student disadvantage and the logic of NCLB". American Journal of Education, 113(3), 367-394.

HEINESEN, E. (2010) Estimating Class-size Effects using Within-school Variation in Subject-specific Classes. The Economic Journal, *120*(545), 737-760.

HOX, J. (2002) "Multilevel Analysis. Techniques and Applications".Mahwah: Lawrence ErlbaumAssociates.

INEE (2013): PISA 2012: Programa para la Evaluación Internacional de los Alumnos. Informe Español. Volumen I: Resultados y contexto, Madrid: Instituto Nacional de Evaluación Educativa.

KROVETZ, M. L. (2007) "Fostering resilience: Expecting all students to use their minds and hearts well". Corwin Press.

LAVY, V. (2010) "Do Differences in School's Instruction Time Explain International Achievement Gaps in Math, Science, and Reading?: Evidence from Developed and Developing Countries". National Bureau of Economic Research, Working Paper 16227.

LEVIN, H. (2010)"A Guiding Framework for Measuring Educational Equity, INES Network for the Collection and the Adjudication of System-Level Descriptive Information on Educational Structures, Policies and Practices".EDU/EDPC/INES/NESLI(2010)6, March 2010.

MARCENARO, O. (2014) "Del lápiz al ordenador: ¿diferentes formas de evaluar las competencias del alumnado?".En INEE (ed.): PISA 2012: Resolución de problemas.

Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MARTIN, A. J. and MARSH, H. W. (2006)"Academic resilience and its psychological and educational correlates: A construct validity approach". Psychology in the Schools, 43(3), 267-281.

MARTINS, L. and VEIGA, P. (2010) "Do inequalities in parents´ education play a important role in PISA students´ mathematics achievement test score disparities?". Economics of Education Review, 29, 1016-1033.

MASLOWSKI, R., ET AL. (2007) "The effect of school autonomy and school internal decentralization on students' reading literacy". School Effectiveness and School Improvement, 18(3), 303-334.

MASTEN, A. S. and COATSWORTH, J. D. (1998) "The development of competence in favorable and unfavorable environments: Lessons from research on successful children". American psychologist, 53(2), 205.

MEDIAVILLA, M. and ESCARDÍBUL, J.O. (2014) "Efecto de las TICs en la adquisición de competencias. Un análisis de género y titularidad de centro para las evaluaciones por ordenador".En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MÉNDEZ, I. (2014)"Factores determinantes del rendimiento en resolución de problemas. España en perspectiva internacional".En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MISLEVY, R. J. (1991) "Randomization-based inference about latent variable from complex samples".Psychometrika 56, Psychometric Society, Greensboro, 177-196.

MISLEVY, R. J. ET AL (1992) "Estimating population characteristics form sparse matrix samples of item responses". Journal of Educational Measurement, 29, 133-161.

NOGUERA, P. A. (2002) "Beyond size: The challenge of high school reform". Educational Leadership, 59, 60-63.

OCDE (2005) "Problem Solving for Tomorrow's World: First Measures of Cross-Curricular Competencies from PISA 2003". Paris: OECD.

OCDE (2008) "Handbook on constructing composite indicators. Methodology and user guide". Paris: Organisation for Economic Cooperation and Development.

OCDE (2009) "PISA 2006 Data analysis manual.SPSS users". Paris: Organisation for Economic Cooperation and Development.

OCDE (2013a) "PISA 2013 Results: Excellence through equity. Giving every student the chance to succeed (Volume II) ".PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201132-en

OECD (2013b) "PISA 2012 Results: Ready to Learn – Students' Engagement, Drive and Self-Beliefs (Volume III) ". PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201170-en

OECD (2014) "Skills for Life: Student Performance in Problem Solving. PISA 2012 Results, vol. VI". Paris: OECD.

PADRON, Y. N. Et al (1999)"Classroom Behavior and Learning Environment Differences Between Resilient and Nonresilient". Journal of Education for Students Placed at Risk, 4(1), 65-82.

RAUDENBUSH, S. Et al (2004)"HLM 6 (Manual) ". Lincolnwood: Scientific Software International.

REID, K. (2013) "Managing School Attendance: Successful Intervention Strategies for Reducing Truancy".Routledge.

ROBINSON, G. E. (1990) "Synthesis of research on effects of class size". Educational Leadership, 47(7), 80-90.

RUMBERGER, R. (2010) "Education and the reproduction of social inequality in the United States: An empirical investigation". Economics of Education Review, 29(2), 246–254.

RUTKOWSKI, L. Et al (2010) "International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting". Educational Researcher, 39 (2), 142-151.

SCHLOTTER, M, SCHWERT, G., WOESSMAN, L. (2011) "Econometric methods for causal evaluation of education policies and practices: a non-technical guide".Education Economics. 19 (2), 109-137.

SIRIN, S. R. (2005) "Socioeconomic status and academic achievement: A meta-analytic review of research". Review of Educational Research, 75(3), 417–453.

TAJALLI, H. and OPHEIM, C. (2004) "Strategies for closing the gap: predicting student performances in economically disadvantaged schools".Educational Research Quarterly, 28(4), 44-54.

TURNER, R. (2006) "El Programa Internacional para la Evaluación de los Alumnos (PISA). Una perspectiva general". Revista de Educación, núm. extraordinario, 45-74.

VAN EWIJK, R. and SLEEGERS, P. (2010) "The effect of peer socioeconomic status onstudentachievement: A meta-analysis".Educational Research Review, 5(2), 134-150.

YANG, Y. and GUSTAFSSON, J.-E.(2004) "Measuring socioeconomic status at individual and collective levels". Educational Research and Evaluation, 10(3), 259–288.

WANG, M.C. ET AL(1994) "Educational resilience in inner cities". En Wang, M.C., Gordon, E.W. (eds.). Educational resilience in inner-city America: Challenges and prospects (pp. 45-72). Hillsdale, NJ: Erlbaum.

WAYMAN, J. C. (2002) "The utility of educational resilience for studying degree attainment in school dropouts". The Journal of Educational Research, 95(3), 167-178.

WILLMS, J. D. (2004) "Reading Achievement in Canada and the United States: Findings from the OECD Programme of International Student Assessment. Final Report.Learning Policy Directorate Strategic Policy and Planning Human Resources and Skills Development".Canada. http://www.hrsdc-rhdcc.gc.ca/sp-ps/arb-dgra

WU, M. and ADAMS, R. J. (2002)"Plausible Values – Why They Are Important". International Objective Measurement Workshop, New Orleans.

3. LEARNING IN MATHS AND PROBLEM SOLVING

Norberto Corral Blanco
Carlos Carleos Artime
Ángela Blanco Fernández
Itziar García Honrado
Ana Belén Ramos Guajardo
Universidad de Oviedo

SUMMARY

In this paper we study the relationship between maths learning and students' experience in solving problems of everyday life. The association between this and the performance in the PISA problem solving is also analyzed. Among the conclusions drawn it should be noted that, when measuring perseverance with the background questionnaire, it was necessary to consider two factors: the first related to overall perseverance and, secondly, to the consistency in responses. On the other hand in the strategies for solving everyday problems, we can differentiate the students who use a reflective-methodical approach from those using an immediate approach; reflective students tend to have higher performance and choose the most effective maths study strategy depending on the learning situation (control-relation-memory). Greater experience in pure maths is associated with better performances, while in applied mathematics the best results are obtained with moderate experience. This difference is explained not by the "pure-applied" dichotomy as much as by the fact that students who usually perform tasks considered as applied mathematics generally get worse results in other PISA tests (Maths, Sciences, and Language).

INTRODUCTION

One of the objectives of the Programme for International Student Assessment in the edition of 2012 (OECD, 2013a and 2013b) focuses on the evaluation of the strategies and attitudes that 15-year-old students have in problem solving. In 2003 this complement to the analysis of performance in maths, reading and sciences was incorporated into the PISA report, but in 2012 in addition to the usual tests printed on paper, new designs of tests were specifically included to be completed using the computer (Computer-Based Assessment - CBA). The possible interrelationship between the results obtained in these areas and problem solving should not be overlooked, so it is important to carry out a joint study in order to identify the connections between them.

In this article we will focus on the maths-problem solving tandem, whose relationship is reflected in many previous studies. For example, according to the National Council of Teachers of Mathematics (NCTM 2000, p.52): "Problem solving involves committing to a task in which the method of its resolution is not known in advance. In order to find a solution, students should look into their knowledge and, through this process, will frequently develop new mathematical knowledge." Therefore, the challenge of dealing with a problem is more demanding than performing a simple exercise; so, throughout the text we will maintain the distinction made between problems and exercises in the PISA survey.

As well as problem solving in the maths field, PISA 2012 CBA includes problems that arise frequently in everyday life that are not related to a specific field of knowledge and that may require more general and complex mental processes (Funke, 2001). According to Schoenfeld (1985, 1992), there are four classical dimensions in relation to problem solving: base knowledge, heuristic processes, metacognition and affective components. Therefore, the challenge posed by problem solving is both emotional and intellectual. As indicated in the report of the OECD (2013a), both student motivation and attitudes towards learning are cognitive factors that influence their performance in a clear way.

The Spanish PISA 2012 report focused on various student attitudes and dispositions towards problem solving, emphasizing interest, extrinsic motivation, perseverance and the facility for dealing with problems. In this study we try to relate the behavior in some of these meta-cognitive/affective variables with the performance in problem solving and maths, as well as the economic-socio-cultural index of the families.

We shall also analyze different heuristics processes of the students when dealing with a specific problem, which will allow us to establish various profiles of problem-solving strategy. These profiles will then be related to the respective performances and the economic-socio-cultural index.

Additionally, in the first section of the study we perform an analysis of the consistency of the responses to some questions from the background questionnaire from PISA 2012. Several sources of inconsistency are analyzed; for example, due to high levels of subjectivity in the response, or inconsistent answers of the students, that were discovered using control questions.

METHODOLOGICAL FRAMEWORK

The data analyzed, corresponding to the PISA 2012 CBA report on problem solving, in which 10,175 people aged fifteen years old participated, contains information collected through background questionnaires and completion of PISA performance tests.

For the first time in the PISA tests, a block of questions was conducted by using computers. The paper "From the pen to the computer: Different ways of assessing student competences?", carried out by Marcenaro et al., and found in this volume, sets out a study on the impact that this change has on the factors usually addressed in the PISA reports.

In general the responses are perceptions or opinions given by students who may have a certain component of subjectivity; for example, for a given person is not always easy to decide whether they do an activity "sometimes" or "frequently", and furthermore this valuation can change greatly from one individual to another.

Although it may be very difficult to determine the frequency with which some activity is done, the use of the category "frequently" can encompass many different situations in the same group; for example, two types of exercises that are done two and four times week, respectively, are tasks that are done "frequently" although the second is practiced twice as much the first. This uncertainty may be important when establishing some conclusions.

All these aspects must be taken into account both in the procedures of analysis and the conclusions drawn.

To get a rough idea of the consistency of the responses of students to questions with a possible relevant subjective component, it was decided to analyze the responses to the question *'to what extent do each of the following statements describe you?'* , from two items of the block which refers to perseverance :

- *When faced with a problem I give up straight away.*
- *I remain interested in the tasks I start.*

The response choices are:

"Very similar to me ", "Fairly similar to me " , " Somewhat similar to me ", " Not very similar to me " , "Not at all similar to me."

Table 3.1 shows the joint distribution of responses, in percentages, and categories that are considered "very consistent", "fairly consistent", "not very or not consistent" marked with the colours green, yellow and red respectively.

The data analyzed for the PISA 2012 CBA

Table 3.1. Percentages of responses to I remain interested and I give up easily

		I remain interested					
	Similar	Very	Fairly	Somewhat	Not very	Not at all	Total
I give up easily	Very	1,8%	1,5%	1,6%	1,1%	,7%	6,6%
	Fairly	1,4%	3,5%	3,4%	2,1%	,7%	11,1%
	Somewhat	2,4%	8,8%	9,7%	4,0%	,5%	25,4 %
	Not very	5,6%	15,0%	11,5%	3,1%	,4%	35,7 %
	Not at all	8,4%	7,6%	3,2%	1,3%	,7%	21,2%
	Total	19,5%	36,4%	29,4%	11,6%	3,0%	100,0 %

The results indicate that 35.9% have "very consistent" responses, 42.7% "fairly consistent", and 21.5%" not very or not consistent".

Although the assignment of labels is always arguable, it seems clear that around 20% of the students mark responses that may be considered uncertain for this type of questions. Also of interest is the 42.7% of "fairly consistent" responses because they point out the difficulty of distinguishing clearly between adjacent categories, eg:, "very" and "fairly".

Another analysis was carried out, of this same type, referring to a block of questions worded as follows: *"In relation to mathematical concepts, to what extent are you familiar with the following terms?"*, which contains three control items, whose supposed mathematical concepts do not really exist. The percentages of the responses to the three items are shown in Table 3.1.

Table 3.2. Degree of familiarity with inexistent mathematical concepts

Item / Degree of familiarity	Never seen it	Seen it once or twice	Seen it several times	Seen it frequently	Know it well, understand the concept
Genuine number (N=6519)	66,0%	16,2%	8,5%	5,1%	4,2%
Subjunctive scale (N=6523)	67,8%	15,5%	8,7%	4,5%	3,5%
Declarative fraction (N=6518)	68,9%	12,4%	8,8%	4,8%	5,1%

It is seen that the distribution of responses is similar for the three items, with the highest frequencies for the consistent answers "I've never seen it" and "I've seen it once or twice", which together exceed 81% of the total. However, the complementary result is not without value, indicating that about 18% of students say they have seen each of these terms at least "several times".

For a more complete picture of the situation, the percentages of those who claimed to have seen any of the above three terms, ie: at least one of the three, and with what degree, were calculated.

The results indicate that almost 50% had never seen them, 19.6% had seen them once or twice and that a significant percentage of students, 30.5%, say they had seen some of these concepts "at least several times".

These two small examples warn us of the need for caution when drawing conclusions.

It would also have been very interesting to have more detailed information about the student's prior experience in preschool education, because in most of the reports this appears as a very important element in relation to performance (see, eg: Corral et al, 2012). This questionnaire only distinguishes their attendance between categories" did not attend", "yes , for a year or less" and "yes , for over a year".

Exploratory data analyses provided the basis for determining the objectives and procedures of the study and to recode some variables.

The variables on which the study focused are the following:

ESCS. Economic-Social-Cultural Status of the Family.

Experience in problem solving.

Perseverance.

Facility

Solving strategies.

Maths learning.

Study strategies.

Frequency and type of exercises done in class.

Frequency and type of problems solved in class.

RPRO. Performance in the Problem Solving in the PISA tests.

The estimations of the parameters associated with the performance such as averages, percentiles, standard errors, etc. were made following the instructions in the manual of data analysis for SPSS users of the PISA program.

The prior treatment of the data was carried out with the PSPP statistical package (GNU Project, 2014, http://gnu.org/software/pspp) using scripts from that manual. In part of exploratory data analysis SPSS 15.0 was used. The estimation of the parameters, the approximation of the standard errors, etc., was made by implementations (ftp://carleos.epv.uniovi.es/pisa) developed by the authors in the R language (R Project, 2014, http:// www.r-project.org) from the methods described in the manual of data analysis for SPSS users of the PISA program.

When statistical inference procedures were used, such as interval estimation and hypothesis checking, it was firstly checked whether the conditions necessary to ensure the validity of the results were verified.

In order to try to smooth out the possible influence that some of the problems of subjectivity already mentioned can have, we have tried to make inferences about, where possible, large groups of students (in general, over one hundred) .

Moreover, to ensure the robustness of the findings, we have used alternative techniques and found that the results were similar. For example, in the regression analysis with ordinal qualitative predictors, two approaches were tested: using ranges/ranks and considering the variable as quantitative in order to adjust a linear term and, sometimes, another quadratic one, or using binary 0 /1 variables to obtain a coefficient for each level.

After checking that, in both approaches, the determination coefficients are almost equal for all cases, and that the meaning of the models was the same, it was decided to include the ranges/ranks model through having a simpler expression.

It should be mentioned that some of the analyses that have been conducted with factors obtained from techniques of dimension reduction, could have been carried out with the corresponding PISA indicator. In this sense, the correlation between the PISA indicator and our principal factor has always been above 90%, but the use of the factors calculated by us allows us to use the second factor and guarantee the linear independence between the two.

RESULTS OF THE STUDY

As pointed out in the introduction, many statistical studies have been carried out to try to explain in a clear and justified way the general behavior of the students who have responded to the background questionnaire from the PISA 2012 CBA study on aspects related to problem solving and maths learning, as well as to other cognitive/affective variables and the economic-social-cultural status of the families, trying to relate it to the performance in Problem Solving (RPRO). The main results of the study are shown below.

Experience in Problem Solving

One of the most important factors which has an influence on learning is the prior knowledge possessed by the learner (Ausubel, 1968). Learning to solve problems not only requires commitment in its resolution, but also intentional acts that enable learning through experience, which plays a critical role in the ability of students when solving problems (Badger et al, 2012).

The research carried out by Heckman et al. (2006, 2010) reveals that educational intervention can change the disposition of the students towards learning, which affects the performance itself.

A study focused on the influence of perseverance, facility and the strategies used by students in the performance in Problem Solving is shown below.

Perseverance

Heckman (2011) suggests that certain personality traits have an influence when explaining the educational level, career outcomes and state of health of the adult population. These include the relevance of perseverance and the capacity for work and sacrifice.

The ability of students to perform at higher levels depends on their beliefs in that, while the ability and talent for certain subjects may help, mastery is achieved only if students work hard and show the necessary perseverance (OECD, 2014). In this sense, it is found that students with low potential but with a high capacity for work and perseverance are more likely to succeed than those who show more talent but low capacity in setting ambitious goals and stay focused on achieving them.

The report tries to measure the perseverance of students from the responses to the questions set out in Table 3.3. A first analysis of the data indicates that 56.9 % of students do not give up when faced with problems, 35.6 % do not usually put off difficult problems, 55.9 % remain interested in tasks they start, 46.6% continue to work on a task until it is perfect and 41.6 % say they do more than what is expected when faced with a problem.

Table 3.3. Percentage of answers to questions from the Spanish PISA report related with perseverance

Questions	Very similar to me	Fairly similar to me	Somewhat similar to me	Not very similar to me	Not at all similar to me
P1. I give up easily when confronted with a problem.	6,6%	11,2%	25,3%	35,7%	21,2%
P2. I put off difficult problems.	12,3%	20,9%	31.2%	24,1%	11,5%
P3. I remain interested in the tasks I start.	19,5%	36,4%	29,5%	11,6%	3,0%
P4. I keep working on tasks until everything is perfect.	17,9	28,7%	28,7%	19,8%	4,9%
P5. When confronted with a problem, I do more than what is expected from me.	16,3%	25,3%	33,1%	20,4%	4,8%

The way questions one and two are formulated with respect to the rest and the inconsistency detected in the methodological framework, recommends that we analyze in detail what factors may explain the association between these items and the origin of the variability in responses. Therefore, a principal components analysis was carried out on the correlation matrix associated with the above five questions and two factors were obtained that explain approximately 70.3% of the overall variability.

The first factor is an indicator of the "overall perseverance" given that it assigns negative scores to students who show a lack of perseverance, while taking positive values for those who do not give up when faced with difficult situations.

The second factor assigns negative scores to those who tend to select, in all questions, responses of the type "similar to me", while taking positive values for those who act in the opposite way, ie: they tend to choose "not similar to me". On the other hand, students with a clearly defined degree of perseverance, whether high or low, occupy the central positions of this variable. Therefore, this factor reflects the "Consistency in Perception of Perseverance".

"Overall Perseverance" shows a growing relationship with both Performance variables as well as with the economic-social-cultural level, as seen in Table 3.4. In this sense we can consult the study "Determinants of Performance in Problem Solving" by Ildefonso Méndez, which is included in this volume.

In Problem Solving, the score is a little over 460 points among those with a low or very low perseverance, and systematically rises to reach an average of 504 points among those with a high level. This conclusion is consistent with that of numerous studies such as that developed by Greene et al. (2004).

Table 3.4. Average score performance by "overall perseverance"

Categories		Very low	Low	Medium	High	Very High
RPRO	Mean	463,97	465,78	472,57	489,78	503,14
(N=6565)	Standard Error	5,73	5,92	5,79	5,47	5,68
ESCS	Mean	-0,314	-0,293	-0,133	-0,181	-0,007
(N=6552)	Standard Error	0,050	0,042	0,050	0,059	0,060

On the other hand, the "Consistency of Perseverance" does not maintain a growing relationship with performance (Table 3.5), as the average scores tend to form a parabolic curve that peaks in the range (60-80], associated with "not similar to me" responses in the first two questions and "similar to me" in the others. An interesting aspect is the difference in behavior between the two classes at either extreme, since people who tend to choose the "similar to me" responses have lower average performance than those who tend to choose the "not similar to me" responses.

Table 3.5. Average scores performance by "consistency in perception of perseverance"

Percentiles		[0, 20]	(20, 40]	(40, 60]	(60, 80]	(80, 100]
RPRO	Mean	456,27	478,19	487,70	494,32	477,01
(N=6568)	Standard Error	6,23	5,54	4,97	5,82	6,35
ESCS	Mean	-0,221	-0.213	-0,144	-0,128	-0,231
(N=6552)	Standard Error	0,047	0,050	0,045	0,051	0.059

These results indicate that by measuring perseverance with the questionnaire it is necessary to apply a correction that takes into account the consistency of the responses. Therefore, when evaluating the importance of perseverance in performance it is necessary to consider both the overall value and consistency. In order to more objectively measure the interest of the consistency we set out two regression models whose only difference is the participation of this variable.

Model 1

The model obtained to predict performance in the Problem Solving in terms of perseverance helps explain 2.5% of the variability and has the following structure:

RPRO = 478.6 + 16.2 Perseverance

where the standard errors of the coefficients are:

SE (Constant) = 4.2; SE (Perseverance) = 2.2.

Model 2

On including the Consistency variable discussed earlier with the quadratic form the percentage of explained variability rises to 4.7% and has the following structure :

RPRO = 486.8 + 14.8 Perseverance + 6.6 Consistency - 8.2 Consistency^2

The standard errors of the coefficients are :

SE (Constant) = 4.3 ; SE (Perseverance) = 2.2;

SE (Consistency) = 2.1 SE (Consistency^2) = 1.1 .

The consistency value at which its contribution to RPRO changes from increasing to decreasing is 0.80, approximately; this score corresponds to the 80th percentile which, as shown in Table 3.5, marks the border between the last two categories. Furthermore, the higher performance of these variables is associated with those who have a high overall perseverance and consistency around 0.80.
On the other hand, the result is conclusive in pointing out that Perseverance and Consistency have the same explanatory power of Performance in Problem Solving.

The Spanish Report of PISA 2012 Vol. I, reflects a significant gap between the performance of those who have repeated a year and those who haven't, which is also related to other important variables such as the economic-socio-cultural level. Although repetition of a course year is not one of the objectives of this study, it seemed interesting to describe in this group other characteristics such as perseverance, which is always lower in the repeaters, apart from those who respond "I do more than what is expected of me". This accumulation of negative circumstances in the group of repeater students gives an idea of the importance of the problem and the difficulty of finding a solution.

Table 3.6. Relationship between Overall Perseverance and repeating course

Perseverance	Non Repeaters	Repeaters
I give up easily when confronted with a problem.	15,5%	24,1%
I put off difficult problems.	34,2%	31,5%
I remain interested in the tasks I start.	58,5%	49,0%
I keep working on tasks until everything is perfect.	49,1%	39,9%
When confronted with a problem, I do more than what is expected from me.	41,8%	48,0%

As for the gender variable, the differences are almost irrelevant, as noted in the Spanish Report 2012 Vol.I.

Another aspect to highlight is that the Consistency of the Perseverance shows no relevant differences with respect to the economic-socio-cultural level or repetition of course years.

Facility in problem solving

The concept studied in this section addresses two of the classic dimensions of problem solving considered by Schoenfeld (1992): emotional and metacognitive. The emotional dimension is the attitude shown by the students in a situation of problem solving and the reactions it provokes in them. Therefore, it encompasses the positive/negative sentiment that the activity, and the actions they feel compelled to follow, inspire in them. The metacognitive dimension refers to knowledge that the students themselves have about their own knowledge, which will lead to a sense of the cognitive demand that it can entail.

Among the five questions included in the questionnaire shown in Table 3.7, the third and fifth refer to the emotional dimension while the other three relate to metacognitive dimension.

Table 3.7: Questions about ability in problema solving

	Very similar to me	Fairly similar to me	Somewhat similar to me	Not very similar to me	Not at all similar to me	Total
I can handle a lot of information.	18,8%	32,9%	34,6%	11,7%	2,0%	100%
I understand things quickly	20,5%	34,3%	29,8%	12,6%	2,8%	100%
I seek explanations for things.	30,0%	35,3%	24,9%	8,2%	1,6%	100%
I can easily link facts together.	24,8%	35,6%	28,1%	9,7%	1,8%	100%
I like solving complex problems.	13,8%	16,3%	26,9%	26,0%	17,0%	100%

The facility or disposition of students to face a new mathematical problem is strongly related with performance, as set out in detail in the PISA 2012 report (PISA 2012, Vol.3, chapter 3).

The results shown in Table 3.8 confirm that the greater the facility in Problem Solving of the students, the higher the scores obtained in the Problem Solving test; this relationship goes in both directions and in no case should it be seen as causal. However, as indicated by PISA 2012, Vol.3, a greater natural ability does not guarantee high performance, since this requires effort and perseverance in the study.

Table 3.8. Average score performance by ability

		Ability in problem solving				
		Very low	Low	Medium	High	Very High
RPRO	Mean	445,15	474,42	490,53	499,34	509,40
	Standard Error	5,83	5,00	5,36	6,52	7,19

Despite how it might sometimes seem, students with greater facility for Problem Solving also show greater perseverance (Table 3.9).

Table 3.9. Means of Perseverance by ability

	Ability in problem solving				
	Very low	**Low**	**Medium**	**High**	**Very High**
Perseverane Mean	-0,481	-0,166	0,074	0,309	0,754
Standard Error	0,026	0,035	0,037	0,051	0,038

On the other hand, there is a clearly growing relationship between the economic-social-cultural family level and the perceived facility for Problem Solving, as clearly shown in Table 3.10.

Table 3.10. Means of ESCS by ability

		Ability in problem solving				
		Very low	**Low**	**Medium**	**High**	**Very High**
ESCS	Mean	-0,439	-0,251	-0,129	-0,035	0,108
	Standard Error	0,038	0,050	0,048	0,056	0,059

The combination of these three analyses highlights the importance of encouraging constant work among students both at home and at school, especially among those belonging to families with lower economic-socio-cultural level. To find out about the behavior of this group of students in more detail we can consult the study "Overcoming barriers: Determinants of Performance in schools and students in an unfavorable environment.", by Cordero et al., which is found in this volume, and which analyzes how to isolate the influence of economic-socio-cultural status (ESCS) in order to study which variables can reduce its impact.

In general it is very difficult to modify the social and cultural characteristics of the family environment and this is why it is so important to act on those aspects that are simpler to change. For example, encouraging perseverance in children from preschool education can increase their involvement in tasks and obtain a real improvement in performance.

Regarding the relationship of facility with students' gender, girls tend to consider that they have less facility than boys in solving problems, especially in the speed of learning and the enjoyment of solving difficult problems (see Table 3.11).

Table 3.11. Percentages of “Very similar to me” or “Fairly similar to me” by gender

	Gender	
	Boys	**Girls**
I can handle a lot of information.	57,5%	45,9%
I understand things quickly	61,6%	52,1%
I seek explanations for things.	64,9%	65,7%
I can easily link facts together.	64,0%	56,9%
I like solving complex problems.	36,7%	23,6%

Strategies in problem solving

In order to try to find out which problem solving procedures are used by the students in the PISA background questionnaire, three problems that frequently occur in daily life were formulated and various compatible strategies for their resolution were proposed. The wording and response options for each problem are shown in Tables 3.12, 3.13 and 3.14 respectively.

Table 3.12. Questions about strategies for solving problems 1

Suppose you have been sending text messages from your mobile phone for several weeks. Today, however, you can't send text messages. You want to try to solve the problem.

What would you do? For each suggestion, tick the option that best suits you.	I would definitely do this	I would probably do this	I would probably not do this	I would definitely not do this	Total
I press every possible button to find out what is wrong I press every button possible to	21,6%	30,7%	25,4%	22,4%	100%
I think about what might have caused the problem and what I can do to solve it.	39,2%	46,9%	11,0%	2,9%	100%
I read the manual	14,1%	21,3%	31,0%	33,6%	100%
I ask a friend for help	32,5%	45,5%	15,3%	6,7%	100%

Table 3.13. Questions about strategies for solving problems 2

Suppose you are planning a trip to the zoo with your brother. You don't know which route to take to get there.

What would you do? For each suggestion, tick the option that best suits you.	I would definitely do this	I would probably do this	I would probably not do this	I would definitel y not do this	Total
I read the leaflet zoo to see if it says how to get there	45,1%	38,7%	10,5%	5,7%	100%
I study a map and work out the best route.	24,7%	35,2%	27,9%	12,2%	100%
I leave it to my brother to worry about how to get there	8,8%	21,3%	37,7%	32,2%	100%
I roughly know where it is, so I suggest we just start driving	23,0%	43,3%	24,2%	9,6%	100%

Table 3.14. Questions about strategies for solving problems 3

Suppose you arrive at the train station. There is a ticket machine that you have never used before. You want to buy a ticket

What would you do? For each suggestion, tick the option that best suits you.	I would definitely do this	I would probably do this	I would probably not do this	I would definitely not do this	Total
I check how similar it is to other ticket machines I have used	33,3%	47,7%	14,4%	4,6%	100%
I try out all the buttons to see what happens	12,7%	21,8%	36,9%	28,6%	100%
I ask someone for help	31,0%	43,9%	18,1%	7,0%	100%
I try to find a ticket office at the station to buy a ticket	33,7%	43,1%	17,4%	5,9%	100%

In general the students usually choose a combination of two or three strategies that vary from problem to problem and, in each case, there is usually one that is clearly in a minority.

The preferences of students in the various problems were:

In the problem of text message (Table 3.12) the most common action is "think about the problem" (Alternative 2), with 86.1% of responses in the "I would definitely do it" or "I would probably do it" while the "read the manual" option is selected by 35.4% for those categories.

In the problem of selecting a route (Table 3.13), 83.8% would choose definitely or probably for "read the leaflet" and only 30.1% choose the easy solution "would leave the problem to the brother".

For the ticket machine problem (Table 3.14), the options "ask for help", "look for a ticket office" and "check whether it is similar to other ticket machines" get very high percentages of the definite or probable option with values above 70%, and the "press the buttons" option is down to 34.5%.

To get an overview of the strategies that students prefer and to facilitate the interpretation of the results we have classified each of the action alternatives as follows:

- **Immediate:** an unthinking response to the problem is given without doing an analysis of the factors involved (pressing all the buttons).
- **Easy:** the answer to the problem is delegated to others (ask for help from a friend).
- **Methodical:** an established work plan is followed (read the manual).
- **Reflective:** the factors involved in the problem are analyzed and an optimal response is attempted (study a map to find the best route).

Subsequently we attempted to identify strategies that best characterize each student by adding the scores assigned to each type. Finally, a principal components analysis was applied to the correlation matrix of the new variables, getting two significant factors which explain, almost equally, 66% of the total variability.

The first factor tends to separate people who combine a large number of strategies from those who prefer to use few alternatives. In this regard it should be highlighted that nearly 99% of students chose at least one of the strategies with the levels "would definitely do it" or "would probably do it". Because of its construction, this factor reflects the degree to which students tend to combine more or less strategies when addressing the problems though regardless of the type of strategies they prefer. Students "would definitely do" or "would probably do" the four options in each of the three problems, with percentages of 12%, 9% and 20%, respectively.

In the variables Performance, Perseverance, Facility and ESCS, the biggest variation from the levels of this factor occurs in the case of Performance, where students who opt to use all strategies tend to have lower scores than the rest (Table 3.15). On the other hand, it appears that those who prefer many strategies tend to be more perseverant and consider that they have greater Facility in problem solving, although in these cases the differences are hardly significant.

This factor, when analyzed in isolation, must be interpreted with great caution because, as already noted, it only refers to the number of strategies without taking into account which ones they are. The real interest of this factor arises when we combine its interpretation with the second, which we shall discuss below.

Table3.15. Relationship between problem solving performance, perseverance, ability and ESCS by the number of strategies

Number of strategies		RPO	Perseverance	Ability	ESCS
Very Low	Mean	482,60	-0,144	-0,035	-0,158
	Standard Error	6,44	0,049	0,041	0,051
Low	Mean	492,13	-0,053	-0,042	-0,163
	Standard Error	5,44	0,037	0,035	0,050
Medium	Mean	479,62	-0,019	-0,054	-0,202
	Standard Error	6,20	0,043	0,036	0,053
High	Mean	480,93	0,081	0,055	-0,152
	Standard Error	5,66	0,041	0,041	0,048
Very High	Mean	460,83	0,140	0,208	-0,271
	Standard Error	6,04	0,037	0,033	0,050

The second factor has a much more interesting meaning, since it sets those who prefer an immediate response, with negative values, against those favoring a Reflective-Methodical strategy. Thanks to this factor we find that the type of strategy that students prefer leads to significant and systematic differences in performance.

Table 3.16. Relationship between problem solving performance, perseverance, ability and ESCS by Reflective-Methodical strategy

Reflective-Methodical Strategy		RPO	Perseverance	Ability	ESCS
Very Low	Mean	450,64	-0,209	-0,211	-0,269
	Standard Error	5,23	0,042	0,038	0,050
Low	Mean	457,37	-0,193	-0,177	-0,299
	Standard Error	6,97	0,036	0,044	0,050
Medium	Mean	473,09	-0,077	-0,011	-0,208
	Standard Error	6,34	0,042	0,034	0,053
High	Mean	492,86	-0,106	0,077	-0,145
	Standard Error	4,80	0,035	0,031	0,057
Very High	Mean	514,10	0,332	0,408	-0,055
	Standard Error	5,28	0,039	0,035	0,052

Both the overall perseverance as well as facility have a growing relationship with the Reflective-Methodical strategy, as can be seen in Table 3.16, as the averages of both variables have a clear upward trend. For example, those who least pause to reflect, ie: prefer a immediate strategy (alternative to Reflective-Methodical in the second factor), have averages of -0.216 in perseverance and -0.215 in facility, rising to 0.341 and 0.435, respectively, for those who mostly opt for a Reflective-Methodical action.

When analyzing overall the experience of students in problem solving in everyday life it can be seen that all the factors considered are clearly interrelated and in turn to performance. For example, students with a higher family economic-social-cultural level tend to have more perseverance, greater perceived facility and use the Reflective-Methodical strategy more frequently than those from a less privileged family status. A more in-depth study of this group of students can be consulted in the study of Cordero, Pedraja and Simancas included in this report (Cordero et al, 2014).

In order to quantify all this set of relationships more clearly a regression model was set up to predict performance using the ESCS and the variables measuring perseverance, facility and their preferred strategy in solving everyday problems.

The model obtained is able to predict 18% of the performance variability (R^2 = 0.18) and shows the following structure:

$$\text{RPRO} = 495.97 + 24.26\ \text{ESCS} + 19.67\ \text{Facility} + 3.89\ \text{Consistency} - 7.40\ \text{Consistency}^2 + 15.68\ \text{Reflective-Methodical} - 8.26\ \text{CStrategies} - 4.96\ \text{CStrategies}^2 .$$

The standard errors of the coefficients of the variables are:

SE (Constant) = 4.06; *SE* (ESCS) = 2.99; *SE* (Facility) = 1.90 ;

SE (Consistency) = 1.89; *SE* (Consistency^2) = 1.04; *SE* (Reflective-Methodical) = 1.78;

SE (CStrategies) = 2.11; SE (CStrategies^2) = 1.38 .

In the model, "CStrategies = Combination of strategies" and "Consistency = Consistency of responses to items of perseverance."

An interesting aspect of this result is that the overall perseverance does not appear in the model, since its coefficient of regression is 1.38 and the standard error 2.39 and, therefore, is not significantly different from zero. This is due to the fact that the information it provides on the performance is contained in the Reflective-Methodical Strategy, Consistency and Facility variables.

Maths learning

The assessment of maths is particularly relevant in the PISA 2012 tests, since it is the area of knowledge that is discussed with the greatest detail and accuracy. Specifically, the assessment of maths competence of students is pursued by analyzing how they use what they have previously learned to solve problems that arise in their real world experiences and build new learning.

According to the model of Biggs (1994), learning results from the interplay of three key elements: intention (motivation) of the learner, the process used (strategy), and the achievements obtained (performance). Therefore, apart from the attitudes to the study, it is also necessary to assess what are the learning strategies used by students and their influence when it comes to understanding and resolving exercises and problems put forward in class.

In this section we will study the following aspects related to the experience that students have in maths learning:

• Learning Strategies

• Frequency and types of maths exercises that they do in class

• Frequency and types of maths problems solved in class

Strategies of maths learning

In this section we will follow the blueprint of the Spanish PISA 2012 report, centered on analyzing the issues focussing on finding out the frequency with which students use different strategies in their daily study.

These strategies are classified into three general patterns:

• **Control** Strategies (CS): the students control the learning process and determine at each moment what they need to know and what they do not know.

• **Reflection-relation** Strategies (RS): the students reflect on what they study and seeks to relate this knowledge to other subjects of real life.

• **Memory** Strategies (MS): the students base their learning, above all, on memorizing the concepts and repeating the exercises in class.

In each of the four situations of Table 3.17 the first response corresponds to control strategies, the second to reflection-relation strategies, and the last one to memory strategies.

Table 3.17. Questions about Mathematical Learning Strategies

Situation	Questions For each group of three items, please choose the item that best describes your approach to mathematics
1	When I study for a mathematics test, I try to define what are the most important parts to learn When I study for a mathematics test, I try to understand new concepts by relating them with things I already know When I study for a mathematics test, I learn as much as I can by heart
2	When I study mathematics, I try to realize which concepts I still have not understood properly When I study mathematics, I think off new ways to get the answer When I study mathematics, I make myself check to see if I remember the work I have already done
3	When I study mathematics, I start working out exactly what I need to learn. When I study mathematics, I try to relate the work with things I have learnt in other subjects. When I study mathematics, I go over some problems so often that I feel as if I could solve them asleep.
4	When I do not understand something in mathematics, I always search for more information to clarify the problem I think about how the mathematics I have learnt can be used in daily life. In order to remember the method for solving a mathematics problem, I go through examples again and again.

In situation 1 we analyze the favourite strategy of students when preparing for an exam. As indicated in Table 3.18, note that the most frequent strategy in the preparation of an exam is to control what is not known (41.9%), although this percentage is similar to that of students who try to relate their knowledge to other subjects or real life (38.2%).

Table 3.18. Percentage of students using different learning mathematics strategies

Situation / Strategy	Control	Relation	Memory	Total
1	41,9%	38,2%	19,9%	100%
2	45,2%	20,5%	34,3%	100%
3	51,4%	19,2%	29,4%	100%
4	22,0%	16,0%	62,0%	100%

The questions from situation 2 do not have a context as clearly defined as the previous block, although in this case, they seem to be directed towards establishing the everyday way of working of the students. A first analysis according to Table 3.18 leads us to conclude that the most common strategy in the day-to-day study is to check/control what subject contents are not well assimilated (45.2%) or to review what has already been studied (34.3%).

Situation 3 has many elements in common with block 2, although in this case the questions seem to be directed towards determining which strategy the student follows to establish the objectives of study. In this case, the most common strategy of students according to the data in Table 3.18 is trying to determine, as accurately as possible, what they need to learn (51.4 %) and only a small percentage (19.2%) tries to relate maths to other subjects.

The items of responses in situation 4 are quite heterogeneous since, for example, while the first sets out how to clarify or resolve a concept that has not been properly understood, the third refers to how to internalize the methods of problem solving. The link we have found between these questions is to determine how the student acts in order to assimilate-secure the concepts or methods: in the first option they seek more information to check/control the problem, in the second they seek some connection to real life and in the third they base their approach on repetition. Clearly the most common strategy shown in Table 3.18 is that of repeating problems or exercises as many times as necessary (62%) and those who try to relate maths to everyday life are clearly in the minority (16%).

As for the difference between boys and girls, the percentages for the four blocks of responses and the three profiles of strategies are shown in Table 3.19. In the exam preparation (block 1) small though significant differences were detected, since girls tend to relate the contents more (40.2%) and use less memory (18.2%), while in the boys these percentages correspond to 36.3% and 21.6 % , respectively.

In scenario 2, related to the day-to-day study, boys resort more to relating to other things (25%) than girls (16 %), but do not bother so much with controlling what they do not properly understand or review what they have already studied.

In relation to situation 3, the most common objective in both sexes is to decide what needs to be learned, with percentages slightly over 50%. Only a small difference can be seen in that the boys choose to relate to other subjects a little more (21%) than girls (17%).

Finally, when faced with the options of situation 4, girls and boys show some noticeable differences since the girls prefer the memory option (67% against 57.1% of boys); however, in the case of looking for the relationship of maths to everyday life, the girls do this with a percentage of 11.9%, while in boys the figure rises to 20%.

Table 3.19. Percentage of students using different learning mathematics strategies by gender

Situation / Strategy	Control Boys - Girls	Relation Boys - Girls	Memory Boys - Girls
1	42,1% 41,6%	36,3% 40,2%	21,6% 18,2%
2	42,6% 47,7%	25,0% 16,0%	32,4% 36,3%
3	21,3% 17,0%	50,6% 52,3%	28,1% 30,7%
4	57,1% 67,0%	20,0% 11,9%	22,9% 21,1%

In the final part of this section we analyze the performance in Problem Solving according to the strategies already discussed (Table 3.17). The results are shown in Table 3. 20.

Table 3.20. Means of problem solving performance by learning strategies and situation

Situation / Strategy		Control	Relation	Memory
1	Mean	471,22	495,18	446,80
	Standard Error	4,86	4,56	5,74
2	Mean	481,45	466,94	473,63
	Standard Error	4,37	5,57	5,01
3	Mean	473,40	481,00	476,05
	Standard Error	4,22	6,01	5,27
4	Mean	478,62	470,73	476,13
	Standard Error	6,39	5,55	4,23

When comparing performance in Problem Solving according to the way of preparing for exams (situation 1), we see that those who opt for reflection-relation of knowledge obtain an average score in their performance of 495.18 points, clearly higher than the average 471.22 points average for those using control strategies. Furthermore, the study of maths based on memory turns out to be very inefficient, with an average of 446,80 points.

In the day-to-day study (scenario 2) some differences also appear, although very small, for the strategy that seems to be more effective in solving problems. The highest score corresponds to those who prioritize by knowing their weaknesses with 481.45 points, while the other two strategies get similar scores around 470 points. This result marks a clear difference between the specific preparation of an examination and daily study, since in the first case the optimal seems to be to try to relate things and the second is controlling what is not properly understood.

Situation 3 practically replicates the results in situation 2 on the scores for each type of strategy, since differences between groups are not statistically significant.

Finally, in situation 4 the differences between the three strategies of maths learning have no significant relevance, although this result could be due mainly to the heterogeneity or uncertainty of the context in which questions are asked.

An interesting aspect is focused on determining whether the problem solving strategies are related to maths learning strategies. To do this the scores of the reflective-methodical factor analyzed in Section 1.3 (which assigned positive values to the Reflective-Methodical profiles, and negative values to immediate profiles) with different maths learning strategies in discussed above (Control-Relation-Memory).

As seen in Table 3.21, in the case of preparing for a test (block 1) the reflective-methodical factor exactly replicates the results obtained with the performance (Table 3.20); ie: the average of the factor is greater for the "Relation" (0.1853) strategy, followed by "Control" (-0.0084) and finally "Memory" (-0.2212).

In the rest of the blocks the association between maths learning strategies and the reflective-methodical factor in Problem Solving does not have such a clear structure but, in general, the factor tends to have higher average in maths strategy which is optimal in relation to performance in problem solving.

Table 3.21. Average score of the Reflective-Methodical strategy by blocks and learning mathematics strategies

Situations		Strategy	
	Control	Relation	Memory
1 Mean	-0,008	0,185	-0,221
Standard Error	0,031	0,048	0,054
2 Mean	0,086	0,085	-0,119
Standard Error	0,031	0,058	0,039
3 Mean	0,036	0,130	-0,096
Standard Error	0,029	0,065	0,040
4 Mean	-0,045	0,091	0,142
Standard Error	0,050	0,066	0,025

The results obtained throughout this section indicate that the maths learning strategies and the strategies followed in everyday life have much in common. Thus, it has been found that, in each of the four cases analyzed, the maths strategy that had higher average score in the reflective-methodical factor also tended to get higher performance in the problem solving.

In colloquial language it could be concluded that students with greater capacity for reflection and working methods tend to study according to the strategies that produce better results.

Exercises performed in class

In the background questionnaire students are asked about how frequently they resolve theoretical and applied maths exercises.

The results shown in Table 3.22 indicate that about 73% of the students do exercises in pure maths (equations - items 5, 7 and 9) "frequently", around 20% "sometimes", 5% "rarely" and 2.5% "never". Applied maths exercises (all other items) are much less frequent since the category "frequently" changes to about 20%, a decrease of about fifty points compared to the same frequency for the other type of exercises, while the percentage of all other categories rises.

Table 3.22. Percentages of how often students encountered the following types of mathematical tasks

	Frequently	Sometimes	Rarely	Never	Total
Working out from a <train timetable> how long it would take to get from one place to another	17,9%	48,3%	25,0%	8,8%	100%
Calculating how much more expensive a computer would be after adding tax	20,8%	47,2%	23,5%	8,5%	100%
Calculating how many square metres of tiles you need to cover a floor	30,0%	44,3%	18,2%	7,5%	100%
Understanding scientific tables presented in a newspaper	10,7%	29,9%	35,4%	24,0%	100%
Solving an equation like 6 x2 + 5 = 29	74,5%	18,1%	4,8%	2,6%	100%
Finding the actual distance between two places on a map with a 1:10,000 scale	18,4%	39,1%	32,4%	10,1%	100%
Solving an equation like 2(x+3) = (x + 3)(x - 3)	73,2%	19,2%	5,0%	2,6%	100%
Calculating the power consumption of an electronic appliance per week	13,4%	35,7%	35,4%	15,5%	100%
Solving an equation like 3x+5=17	73,0%	18,9%	5,7%	2,5%	100%

When comparing the performance results in Problem Solving by the frequency with which the different types of exercises are performed clearly shows a systematic and very important difference between pure and applied maths exercises (Table 3.23).

By studying the performance according to the frequency with which the pure maths tasks are performed, it can be seen that highest scores always correspond to the category "frequently", with an average around 490 points, which drops clearly for the rest of the frequencies.

In applied maths the situation changes completely, because now the highest scores are obtained from students who do this type of tasks "sometimes" or "rarely", averaging about 480 points, while for those who do them "frequently" the average is usually around 460 points. The result corresponding to those who never do them has a large variability, with values between 440 and 480 points.

Table 3.23. Average score performance in Problem solving by the frecuency of solving mathematical tasks

		Frequently	Sometimes	Rarely	Never
Working out from a <train timetable> how long it would take to get from one place to another	Mean	455,2	474,3	493,4	469,3
	Standard Error	6,6	4,4	4,9	8,9
Calculating how much more expensive a computer would be after adding tax	Mean	465,8	468,7	482,5	458,9
	Standard Error	5,5	4,3	6,0	9,5
Calculating how many square metres of tiles you need to cover a floor	Mean	477,5	482,5	468,2	442,4
	Standard Error	4,3	4,9	6,7	8,9
Understanding scientific tables presented in a newspaper	Mean	453,9	469,7	483,1	480,7
	Standard Error	8,0	5,2	4,5	6,4
Solving an equation like 6 x2 + 5 = 29	Mean	490,1	444,8	405,3	389,8
	Standard Error	3,9	5,6	11,7	16,2
Finding the actual distance between two places on a map with a 1:10,000 scale	Mean	452,1	476,6	490,4	465,4
	Standard Error	5,3	4,6	4,9	8,6
Solving an equation like 2(x+3) = (x + 3)(x - 3)	Mean	489,1	450,7	407,3	397,4
	Standard Error	4,1	6,1	9,9	16,0
Calculating the power consumption of an electronic appliance per week	Mean	451,2	476,3	486,3	467,8
	Standard Error	6,8	4,4	4,9	7,3
Solving an equation like 3x+5=17	Mean	486,3	456,4	451,3	393,3
	Standard Error	4,0	6,0	11,0	16,1

In order to describe all this information clearly and concisely the variables "Experience of Pure Maths" and "Experience of Applied Maths" have been used for summarizing the frequencies of each type of tasks. Both variables were divided into four categories that roughly correspond to those of the questionnaire, ie: "never", "rarely", "sometimes" and "frequently".

The averages of performance according to these categories clearly summarize what was mentioned when describing the various tasks performed in class (Tables 3.24 and 3.25).

Table 3.24. Average score performance in Problem solving by the frecuency of solving pure mathematical tasks

		Experience on Pure Mathematical Tasks			
		Frequently	Sometimes	Rarely	Never
RPRO	Mean	490,17	449,06	409,32	395,47
	Standard Error	3,98	6,04	9,28	16,97

Table 3.25. Average score performance in Problem solving by the frecuency of solving applied mathematical tasks

		Experiencia en matemática aplicada Experience on Applied Mathematical Tasks			
		Frequently	Sometimes	Rarely	Never
RPRO	Mean	455,95	479,22	486,75	470,57
	Standard Error	5,29	3,94	6,89	8,15

A more detailed study of this situation, which analyzes the experience in applied maths and pure maths together in relation to the performance, points to a surprising result, since doing applied maths exercises "frequently" appears to be associated with a drop in performance, regardless of what happens with pure maths exercises (Table 3.26).

Table 3.26. Relation between pure and applied mathematical tasks with performance

Experience on Pure Mathematical Tasks		Experience on Applied Mathematical Tasks			
		Frequently	Sometimes	Rarely	Never
Never	Mean	362,13	395,72	416,92	392,02
	Standard Error	85,37	39,31	26,64	18,71
Rarely	Mean	366,25	303,74	324,91	319,11
	Standard Error	16,27	14,61	13,79	23,89
Sometimes	Mean	404,88	448,54	461,93	476,74
	Standard Error	14,47	6,32	10,13	12,65
Frequently	Mean	467,56	493,45	503,97	504,41
	Standard Error	5,69	3,96	7,56	9,66

Note: The standard errors in the "Never" category of experience in pure maths are due to the fact that there are few people who fulfill this condition.

In order to go into some more depth in this analysis we have studied the economic-social-cultural status of students' families, according to the frequency of the tasks of pure and applied maths.

The results in Table 3.27 indicate a slight difference of the ESCS with respect to experience in applied maths, observed mainly in the category "never", which has a weight of 10% in the sample.

Table 3.27. Means of ESCS according to the experience in solving applied mathematical tasks

	Experience on Applied Mathematical Tasks			
	Frequently	Sometimes	Rarely	Never
ECSC Mean	-,1475	-,1613	-,1882	-,2356
Standard Error	,051	,035	,056	,074

With regard to the experience in pure maths, the results in Table 3.28 reflect that the differences between families are more pronounced and a noticeable increase in the economic-social-cultural status on going from "observed rarely" to "frequently". This trend is not maintained in the category "never", although it is hardly representative both because of its weight in the sample, about 2.5%, and the chance that a student of fifteen years old has never seen equations of first or second degree is not very consistent.

Table 3.28. Means of ESCS according to the experience in solving pure mathematical tasks

	Experience on Pure Mathematical Tasks			
	Frequently	**Sometimes**	**Rarely**	**Never**
ECSC Mean	-,1009	-,2975	-,5452	-,3898
Standard Error	,036	,048	,083	,097

To determine whether these differences observed in the social status of families may explain the observed differences in performance, a regression model was set out, which used the variables of Experience of Pure Maths, Experience of Applied Maths and Economic-Social-Cultural Status.

The model obtained, explaining 14.2% of the performance variability ($R^2 = 0.142$), is as follows:

RPO = 472.67 + 25.35 ESCS + 33.96 Pure Maths - 15.42 Applied Maths

The standard errors of the coefficients are:

ST (Constant) = 4.05; *ST* (ESCS) = 2.96;

ST (Pure Maths) = 2.74; *ST* (Applied Maths) = 2.77.

The conclusion from this model is that the ESCS and the frequency of performing pure maths exercises are positively associated with performance, while for applied maths the most efficient seems to be performing these tasks only "sometimes" or "rarely". The effect of "never" is almost irrelevant and can be collected using a quadratic model in "Experience in Applied Maths", but the regression coefficient of this variable is not significantly different from zero.

Types of problems

The background questionnaire of PISA 2012 asks about the frequency of resolution of four different types of problems that are set in class:

- Algebraic.
- Procedural.
- Pure mathematical reasoning.
- Applied mathematical reasoning.

In addition, at the beginning of the questions two specific examples of each type are set and students are asked how frequently they do them, both in class and in school exams. See Tables 3.29, 3.30, 3.31 and 3.32.

Table 3.29. Questions about algebraic problems

There are some problems below. Each one requires you to understand a problem and perform the appropriate calculations. The problems usually refer to practical situations, but the numbers, people and places mentioned are made up. All the information you need is given. Here are two examples:

1) <Ann> is two years older than <Betty> and <Betty> is four times as old as <Sam>. When <Betty> is 30, how old is <Sam>?
2) Mr <Smith> bought a television and a bed. The television cost <$625> but he got a 10% discount. The bed cost <$200>. He paid <$20> for delivery. How much money did Mr <Smith> spend?

Table 3. 30. Questions about procedural problems

Below there are examples of another set of mathematical skills.

1) Solve 2x + 3 = 7.
2) Find the volume of a box with sides 3m, 4m and 5m.

Table 3. 31. Questions about pure mathematical problems

In the next type of problem, you have to use mathematical knowledge and draw conclusions. There is no practical application provided. Here are two examples.

1) Here you need to use geometrical theorems:
   Determine the height of the pyramid.
2) Here you have to know what a prime number is:
   If n is any number: can (n+1)2 be a prime number?

Table 3. 32. Questions about applied mathematical problems

In this type of problems, you have to apply suitable mathematical knowledge to find a useful answer to a problem that arises in daily life or work. The data and information are about real situations. Here are two examples.

Example 1:
A TV reporter says “This graph shows that there is a huge increase in the number of robberies from 1998 to 1999”.
Do you consider the reporter’s statement to be a reasonable interpretation of the graph? Give an explanation to support your answer. (See the graph)

Example 2:
For years the relationship between a person’s recommended maximum heart rate and the person’s age was described by the following formula:
Recommended maximum heart rate = 220 - age
Recent research showed that this formula should be modified slightly. The new formula is:
Recommended maximum heart rate = 208 - (0.7 Ã— age)
From which age onwards does the recommended maximum heart rate increase as a result of the introduction of the new formula? Show your work.

The problems that are most frequently done in class are procedural, as 72.5% of the students do them "frequently", followed by algebraic, with 58.1%, while these percentages are significantly reduced in the mathematical reasoning problems, whether pure or applied, with values of 35.4% and 23.3% respectively (Table 3.33).

Table 3.33. Percentages of solving different types of mathematical problems frecuency

	Types of Problems			
	Algebraic	Procedural	Pure mathematical reasoning	Applied mathematical reasoning
Frecuently	58,0%	72,4%	35,4%	23,4%
Sometimes	36,6%	23,1%	44,3%	48,3%
Rarely	4,0%	3,5%	16,3%	23,4%
Never	1,4%	1,0%	4,0%	4,9%
Total	100%	100%	100%	100%

By studying the relationship between the frequency of resolution of each type of problems and performance (see Table 3.34), it is seen that in the case of algebraic and procedural problems students decrease their performance when reducing the frequency with which they perform this type of problems, since they go from an average of around 480 points for the category "frequently" to about 400 points for the category "never".

In algebraic problems the drop in scores occurs gradually, about 20 points in the first three categories and 50 points in the next.

In the procedural problems the average performance associated with the categories "sometimes" and "rarely" is almost identical and is 30 points below the average value of the category "frequently".

Performance differences in the categories of pure mathematical reasoning are much less pronounced than in the ones above, and have little relevance except in the option "never".

Table 3.34. Average score performance in problem solving by types of mathematical problems

	Types of mathematical problems solved in the classroom							
	Algebraic		Procedural		Pure mathematical reasoning		Applied mathematical reasoning	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
Frecuently	487,0	3,96	484,7	4,09	481,6	4,87	461,0	5,07
Sometimes	465,6	5,35	455,9	6,20	47564	4,85	477,0	4,25
Rarely	449,5	12,80	454,8	15,75	476,2	6,40	490,3	5,66
Never	392,0	16,78	402,9	23,53	440,7	10,99	470,8	11,75

In the case of applied mathematical reasoning, students who get higher performance do this kind of problems "rarely", getting a truly surprising result since they have a higher average score in the category "never", than when they do them "frequently".

In order to determine an overview of the relationship of the four types of problems with Problem Solving, a regression model is used in which the economic-social-cultural status is included to eliminate differences that might exist among students owing to this factor .

The model obtained accounts for about a 10% of the performance and has the following structure:

RPRO = 501.39 + 26.13 ESCS + 17.27 Algebraic + 14.13 Procedural - 12.61 * Applied Reas.

The standard errors of the coefficients are:

ST (Constant) = 7.29; ST (ESCS) = 3.07; ST (Algebraic) = 3.42

ST (Procedures) = 3.95; ST (Applied Reas.) = 2.37

The problems of pure mathematical reasoning do not intervene in the model because their regression coefficient is worth 1.14 and the standard error is 2.48, which turns out to be not significantly different from zero. Therefore, the general conclusion is the same as with the exercises in class, discussed above in Section 2.2.

What happens with applied maths?

The aim of this section is to find some explanation for the unexpected performance that was associated with the frequency in performing exercises or problems of applied maths.

From a logical point of view, it makes little sense that students who solve a lot of algebraic or procedural problems reduce their score in Problem Solving because they do a lot of applied maths problems.

For this reason it may be interesting to address some aspects to clarify this question. Specifically, we will focus on the following points:

a) Analyze in detail the questionnaire items.

b) Identify the "academic" profile of the students that say they do many applied maths problems and compare it with the rest.

Analysis of the questionnaire items

With regards to the analysis of the items from the background questionnaire, which refer to mathematics, we will study two specific situations in which a very clear relationship between an item and a type set of problems is established.

Situation 1

This situation relates to item 2 of maths exercises and the second of the algebraic problems with a statement applied. These are the following:

Item 2: Calculate how much the price of a computer increases when adding the taxes.

Problem: Mr. Smith has purchased a TV and a bed. The TV cost 625 €, but he has got a 10% discount. The bed costs 200 €. He has paid 20€ on the transport home. How much money has Mr. Smith spent?

The problem is somewhat more complex because of the length of the statement, but the greatest mathematical difficulty of the problem is in applying a tax or getting a discount. From this point of view both tasks can be considered Applied Maths problems or exercises of similar difficulty.

When considering the performance along with the item and the problem response, a regression model can be constructed that explains 9.7% of the overall variability (R^2 = 0.097) and that has the following expression:

RPRO = 464.46 + 27.15 ESCS +21.59 Algebraic - 42.97 Prices - 8.15 Prices2 ,

The standard errors of the coefficients are :

SE (Constant) = 11.00; SE (ESCS) = 2.97; SE (Algebraic) = 3.44;

SE (prices) = 9.91; SE (Prices2) = 2.01 ;

where prices represents the frequency of completion of item 2.

Note: The quadratic term of the prices was included in the model because its relationship to the Performance resembles a parabolic curve.

By studying the above model it is clearly seen that the increase of ESCS or frequency of algebra problems improve performance, while doing such exercises (items) frequently reduce it.

From this perspective it seems that the algebraic problems related to a real problem are not identified by students as problems of Applied Maths, but in reality they are working on the same competences.

Situation 2

This situation refers to item 9 of the maths exercises and the second problem of applied mathematical reasoning, these being:

Item 9: Solving an equation like the following: $3x + 5 = 17$.

Problem: For years, the relationship between the maximum recommended heart rate for a person and his/her age was expressed by the following formula :

Recommended maximum heart rate = 220-age

Recent research has shown that this formula had to be modified slightly. The new formula is as follows :

Recommended maximum heart rate = 220 - (0.7 x age)
From what age does the maximum recommended heart rate increase as a consequence of the introduction of the new formula? Justify your answer.

The item is a purely procedural exercise, as it is about solving an equation of the first degree. However, the applied problem is more complicated as it is necessary to understand the statement, to deal with the equation of a straight line and also the final solution involves solving the first degree equation $220 - x = 208 - 0.7x$, which is equivalent to solving $0.3x + 208 = 220$.

Therefore, working with this type of problems would involve understanding the statement (reading comprehension), setting out the model to be solved (formulating the equation of the first degree) and apply the procedural skills (solving the equation) .

Since the students are asked to read the statement of the problems but not to solve it, you would not expect a detailed analysis of what it would really involve, though solving an applied maths problem frequently requires all of the elements mentioned above.

From a strictly mathematical perspective it is very difficult to understand that the fact of resolving these problems "frequently" lowers the performance, while the procedural problems increase it, as reflected in the following regression model, which explains 11,1% of the performance variability :

RPRO = 497.62 + 26.41 ESCS - 9.15 Applied Reas. + 26.03 Equation

The standard errors of the coefficients are :

SE (Constant) = 6.45; SE (ESCS) = 3.08; SE (Applied Reas.) = 2.30; SE (Equation) = 2.81

where Applied Reas. = frequency of problem solving of Applied Mathematical Reasoning

Equation = frequency of resolution of item 9 (linear equation).

On the other hand, the two statements that are put forward in the applied problems are very different in nature; in the first the difficulty is in the statement because from the point of view of procedure it only involves a subtraction and a quotient, while in the second the need to know much more complex mathematical procedures is added to the difficulty of the statement.

In order to determine the "academic" profile of the students that claim to do more exercises in applied maths or reasoning problems of applied maths, the results of other PISA tests such as Reading and Science were calculated, in addition to Maths and Maths with computers. The corresponding results are detailed in the following sections.

Academic profile of the students

In regard to performance in the various PISA tests regarding experience with maths exercises, the results are completely consistent with what is already known, since they replicate the structure of performance found in the Problem Solving (see Table 3.35). Higher scores are obtained among students claiming to do Applied Maths homework "sometimes" or "rarely", regardless of the type of matter.

Table 3.35. Average score performances by the experience with applied mathematics

Experience with applied mathematics		PV Mathematics	PV Lecture	PV Science	PV RPRO	PV CBA MATHS
Never	Mean	475,92	480,65	481,79	470,56	461,06
	Standard Error	5,71	6,40	5,29	8,55	5,42
Rarely	Mean	493,63	499,73	504,73	486,75	478,02
	Standard Error	4,30	4,53	4,10	4,89	4,61
Sometimes	Mean	488,05	491,24	503,71	479,22	479,41
	Standard Error	2,89	3,23	2,64	3,95	3,66
Frequently	Mean	472,82	472,35	490,79	455,95	472,96
	Standard Error	4,87	4,53	4,35	5,29	4,52
Total	Mean	482,95	485,57	494,87	476,77	475,08
	Standard Error	2,43	2,65	2,37	4,10	3,17

When the above analysis is performed by taking into account the experience in pure maths, the results are identical in structure to those already obtained, since the average values in all PISA tests grow by increasing the frequency with which these tasks are performed (see Table 3.36).

Table 3.36. Average score performances by the experience with pure mathematics problems

Experience with pure mathematics		PV Mathematics	PV Lecture	PV Science	PV RPRO	PV CBA MATHS
Never	Mean	438,51	424,00	437,81	395,47	432,58
	Standard Error	13,15	12,82	12,54	16,97	10,88
Rarely	Mean	438,56	423,34	448,28	409,32	440,34
	Standard Error	6,81	6,76	6,44	9,28	6,90
Sometimes	Mean	463,01	459,49	476,73	449,06	456,51
	Standard Error	4,34	5,01	4,34	6,04	4,75
Frequently	Mean	495,89	503,15	511,12	490,17	485,40
	Standard Error	2,84	3,06	2,58	3,98	3,33
Total	Mean	482,95	485,57	494,87	476,77	475,08
	Standard Error	2,43	2,65	2,37	4,10	3,17

The distribution of the averages of the PISA test results by taking into account the classification of the maths problems as algebraic, procedural, pure mathematical reasoning and applied mathematical reasoning again replicates what was seen in the case of Problem Solving. In the first three types of problems the increase in the frequency of resolution is associated with an increase in performance, while for the latter type of problems the relationship changes completely (Tables 3.37, 3.38, 3.39 and 3.40).

Table 3.37. Performance according to solving algebraic problems

Algebraic problems		PV Mathematics	PV Lecture	PV Science	PV RPRO	PV CBA MATHS
Never	Mean	428,28	418,39	438,23	392,01	417,52
	Standard Error	13,46	14,47	13,30	16,78	15,42
Rarely	Mean	475,84	477,89	484,50	449,51	465,01
	Standard Error	8,71	10,75	7,94	12,80	8,34
Sometimes	Mean	475,35	473,07	488,36	465,56	467,67
	Standard Error	3,52	3,91	3,85	5,35	3,72
Frequently	Mean	493,60	501,34	508,79	486,99	484,69
	Standard Error	3,00	3,24	2,62	3,96	3,29
Total	Mean	482,95	485,57	494,87	476,77	475,08
	Standard Error	2,43	2,65	2,37	4,10	3,17

Table 3.38. Performance according to solving procedural problems

Procedural problems		PV Mathematics	PV Lecture	PV Science	PV RPRO	PV CBA MATHS
Never	Mean	429,98	412,43	431,56	402,95	413,15
	Standard Error	17,42	17,77	16,32	23,53	18,37
Rarely	Mean	467,28	454,53	481,87	454,81	465,19
	Standard Error	10,34	10,16	9,37	15,75	11,30
Sometimes	Mean	469,08	465,50	484,50	455,92	461,07
	Standard Error	4,41	4,82	5,24	6,20	4,25
Frequently	Mean	491,90	498,85	505,78	484,69	483,00
	Standard Error	2,75	2,93	2,52	4,09	3,46
Total	Mean	482,95	485,57	494,87	476,77	475,08
	Standard Error	2,43	2,65	2,37	4,10	3,17

Table 3. 39: Performance according to solving pure mathematical problems

Pure mathematical reasoning Problems		PV Mathematics	PV Lecture	PV Science	PV RPRO	PV CBA MATHS
Never	Mean	452,99	455,45	469,08	440,70	444,99
	Standard Error	7,04	7,10	6,98	10,99	7,49
Rarely	Mean	479,13	483,12	495,42	476,20	472,14
	Standard Error	4,50	5,00	4,95	6,40	4,23
Sometimes	Mean	485,43	486,90	498,38	475,64	475,81
	Standard Error	3,12	3,57	3,12	4,85	3,68
Frequently	Mean	492,28	498,63	506,56	481,59	483,65
	Standard Error	3,99	3,85	3,39	4,87	4,36
Total	Mean	482,95	485,57	494,87	476,77	475,08
	Standard Error	2,43	2,65	2,37	4,10	3,17

Table 3.40: Performance according to solving applied mathematical problems

Applied mathematical reasoning problems		PV Mathematics	PV Lecture	PV Science	PV RPRO	PV CBA MATHS
Never	Mean	475,82	488,95	491,87	470,82	471,04
	Standard Error	8,35	8,99	7,68	11,75	7,65
Rarely	Mean	499,91	503,02	512,38	490,29	484,99
	Standard Error	3,78	4,44	3,89	5,66	3,88
Sometimes	Mean	484,72	487,44	498,31	477,01	476,37
	Standard Error	3,25	3,36	3,02	4,25	3,63
Frequently	Mean	473,37	477,35	489,71	461,03	469,27
	Standard Error	3,32	3,70	3,06	5,07	4,49
Total	Mean Standard Error	482,95 2,43	485,57 2,65	494,87 2,37	476,77 4,10	475,08 3,17

These last analyses indicate that the cause of the negative relationship between applied maths problem solving and performance in Problem Solving is not due to these exercises in themselves. In reality, the central theme of this strange behavior is that students who most frequently solve these exercises are those with lower performance in all tests.

On the other hand, in addition to the Problem Solving performance, there are other aspects that must be taken into account in the development of students. Specifically, in the background questionnaire are included such important metacognitive aspects as perseverance, already analyzed in detail in section 1.1, facility of problem solving, described in section 1.2, or the interest, self-efficacy and self-esteem in maths .

In Tables 3.41 and 3.42 the results of this analysis are shown. As can be seen, the frequency of performing both types of tasks has a positive relationship with all cognitive variables identified. Regarding pure maths tasks, the results complement those already obtained in terms of improved performance in maths and problem solving according to the resolution of types of algebraic, procedural and pure mathematical problems analyzed above (Tables 3.29, 30 and 3.31).

Table 3.41. Cognitive variables according to the frecuency of pure mathematical tasks

		Interest	Self-Efficacy	Self-esteem	Overall Perseverance	Ability
Never	Mean	-,8376	-,7767	-,2508	-,380	-,627
	Standard Error	,152	,312	,152	,258	,188
Rarely	Mean	-,2203	-,2514	-,1465	-,247	-,182
	Standard Error	,069	,084	,098	,080	,789
Sometimes	Mean	-,2437	-,1580	-,2050	-,074	-,133
	Standard Error	,059	,052	,062	,045	,044
Frequently	Mean	-,0317	,1941	-,0456	,187	,076
	Standard Error	,035	,024	,034	,028	,029

Table 3.42. Cognitive variables according to the frecuency of applied mathematical tasks

		Interest	Self-Efficacy	Self-esteem	Overall Perseverance	Ability
Never	Mean	-,3723	-,2451	-,1722	-,167	-,238
	Standard Error	,077	,091	,067	,091	,106
Rarely	Mean	-,1961	-,0377	-,0457	,015	-,084
	Standard Error	,047	,048	,055	,046	,046
Sometimes	Mean	-,1080	,0555	-,0642	,089	-,021
	Standard Error	,042	,032	,038	,039	,027
Frequently	Mean	,0403	,2620	-,1226	,224	,166
	Standard Error	,052	,048	,064	,047	,047

As regards the tasks of applied maths, they play an important role in education, as they are positively related to almost all metacognitive aspects mentioned. For example, Table 3.42 shows that by increasing the frequency of performing this type of tasks, this achieves an increase in the average values of both interest and self-efficacy in maths, as well as perseverance and facility.

Conclusions

A multitude of factors intervene in the learning processes, most of them interrelated, and which are associated with performance but which never determine it.

"Academic" learning is a process that begins in the early years of the peoples' lives and it has a cumulative character. That is why it is so important for the person to acquire good study habits in aspects such as perseverance, learning strategies, self-confidence, etc., while avoiding leaving gaps in knowledge.

There is a relatively high percentage of students, about 30%, who say they have seen "at least several times" some mathematical concepts contained in the background questionnaire as control questions, but that do not really exist. It was also observed that about 20% of the students responded in a "not very or not" consistent way to some questions about perseverance and formulated so that the same response category has the opposite meaning.

Perseverance is clearly related to performance and, in general, the people who have this quality more developed get better results, even when starting out from unfavorable conditions. This study has found that when perseverance is measured by the survey questions, the consistency of responses should also be taken into account since the coherence of students in their responses is associated with higher performance.

The self-perception of the facility of problem solving is also positively related to performance. On the other hand, people who perceived themselves as having more facility also tended to be more perseverant. This result again shows the overlap/interdependence between the factors related to the performance, and the fact that improving the behavior of one of them can produce an increase in performance that goes beyond this one-off improvement.

Students tend to combine different strategies for solving problems of everyday life, while granting them different priorities depending on the situation employed. In addition it was found that people who tend to deal with problems in daily life based on a reflective-methodical response, combined with some other strategy, tend to get the best results in Performance.

Another interesting result is that the group of students who opt to implement fewer strategies tend to have a lower performance than those who choose to try multiple options.

The most common maths exercises in classes are procedural (solving equations), since over 70% of students do them "frequently". Applied maths exercises are far less usual, since they are only performed "frequently" by about 20%.

The relationship of performance to the two types of maths exercises (pure and applied) is completely different. In the case of pure maths, the performance increases with the frequency of resolution of this type of exercises. For applied maths, the best performance is achieved when these tasks are done "rarely" or "sometimes". When performance is studied taking into account the classification of problems such as: algebraic with an applied statement, procedural, pure mathematical reasoning and applied mathematical reasoning, the results obtained in the case of the exercises are corroborated.

Performance has a positive relationship with the frequency of algebraic or procedural problem solving. In the case of problems of pure reasoning, the same trend continues, but the differences are much less pronounced.

By analyzing the problems of applied mathematical reasoning a very striking result is obtained, since the performance tends to decrease when such problems are "frequently" resolved, regardless of the frequency with which other types of problems are done. Moreover, the best performance is obtained when the applied mathematical reasoning problems are done "rarely".

When taking into account the relationships between exercises and problems, apparently contradictory results are found. For example, in calculating the amount of a purchase using taxes or discounts, it was found that, when considered as an exercise in applied maths, those who do it "frequently" have average scores of 465.8, lower than those who do it "rarely", with an average of 482.5; but when considered as an applied algebraic problem the situation is reversed since among those who do it "frequently" the average is 487.0 and drops to 449.5 for "rarely". The same happened with the resolution of a first degree equation, as considered in the questionnaire as an exercise in pure maths or as a applied mathematical reasoning problem.

The explanation that seems more credible for these apparent contradictions lies not so much with the pure maths-applied maths dichotomy, as in the student characteristics that tend to do more frequently those tasks considered to be applied maths. After analyzing the scores of these students in other PISA tests, such as science or reading, it was found that the structure of the performances is the same in all of them, ie: such students have lower scores not only in problem solving or maths but also in reading and sciences.

It is worth highlighting that the so-called applied maths exercises in the questionnaire involve solving processes and more simple mathematical concepts (basic arithmetic and geometry) than so-called pure maths exercises (quadratic/second degree equations). One might think that the "applied" problems are seen more frequently by students who may have more difficulty with advanced mathematical concepts. A clear conclusion can not be reached since the category of "frequently" may be including very distinct actual frequencies.

On the other hand, applied maths, which is more frequent among students with low performance, has some very important features that must be taken into account; for example, it appears to be associated with an improvement in self-efficacy or interest in maths, as well as the facility and perseverance for solving problems.

REFERENCES

AUSUBEL, D.P. (1968). "Educational Psychology: A Cognitive View." Holt, Rinehart & Winston.

BADGER, M.S., SANGWIN, C.J., HAWKES, T.O. (2012). "Teaching Problem-Solving in Undergraduate Mathematics", University of Birmingham.

BIGGS, J.B. (1994). "Student learning research and theory: where do we currently stand?". En G. Gibbs (Ed)., *Improving student learning: Theory and practice*, pp.1-19. Oxford: The Oxford Centre for Staff Development.

CORDERO, J.M., PEDRAJA, F., SIMANCAS, R. (2014) "Superando las barreras: factores determinantes del rendimiento en escuelas y estudiantes con un entorno desfavorable", *Informe PISA 2012 CBA.*

CORRAL BLANCO, N., ZURBANO FERNÁNDEZ, E., BLANCO FERNÁNDEZ, A., GARCÍA HONRADO, I., RAMOS GUAJARDO, A. (2012). "Estructura del entorno educativo familiar: su influencia sobre el rendimiento y el rendimiento diferencial", *Informe PIRLS-TIMSS 2011, Estudio Internacional de progreso en comprensión lectora, matemáticas y ciencias, Volumen 2: Informe español. Análisis secundario*, pp. 10-42.

FUNKE, J. (2001). "Dynamic systems as tools for analysing human judgment", *Thinking & Reasoning* 7 (1), pp. 69–89.

GREENE, B.A., MILLER, R.B., CROWSON, H.M., AKEY, K.L. (2004). "Predicting high school students' engagement and achievement: Contributions of classroom perceptions and motivation", *Contemporary Educational Psychology* 29, pp. 462-482.

GREIFF, S., HOLT, D.V., FUNKE, J. (2013). "Perspectives on Problem Solving in Educational Assessment: Analytical, Interactive, and Collaborative Problem Solving", *Journal of Problem Solving* 5 (2), pp. 71-91.

HECKMAN, J.J., STIXRUD J., URZUA, S. (2006). "The effects of cognitive and non-cognitive abilities on labor market outcomes and social behaviour", *Journal of Labor Economics*, 24 (3), pp. 411-482.

HECKMAN, J.J., MOON, S.H., PINTO, R., SAVELYEV, P.A., YAVITZ, A. (2010). "The rate of return to the HighScope Perry Preschool Program", *Journal of Public Economics* 94, pp. 114-28.

HECKMAN, J.J. (2011) "Integrating Personality Psichology into Economics", *NBER Working Paper* No. 17378.

MORALES VALLEJO, P. (2011). "Cuestionarios y escalas", *Universidad Pontificia*

Comillas, Madrid. Disponible en
http://www.upcomillas.es/personal/peter/otrosdocumentos/CuestionariosyEscalas.doc

NTCM (2000) "Principles and Standards for School Mathematics" *National Council of Teachers of Mathematics (NCTM),* Reston.

OECD (2013a) "PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy", París.

OECD (2013b) "PISA 2012 Results: What Makes a School Successful? Resources, Policies and Practices. Vol. IV", París.

OECD (2014) "Do students have the drive to succeed?" PISA In Focus, 37.

SCHOENFELD, A. H. (1985). Mathematical problem solving. San Diego, California: Academic Press.

SCHOENFELD, A. H. (1992). Learning to think mathematically: Problem solving, metacognition and sense making in mathematics. En *Handbook for Research on Mathematics Teaching and Learning*. Ed. Grouws, 334-370.New York,:Macmillan.

4. From the pencil to the computer: Different ways of assessing students' competences?

Óscar D. Marcenaro Gutiérrez

Universidad de Málaga

Summary

The implementation of computer-based assessment (CBA) in PISA and the OECD's intention of using this type of assessment mode for the tests which are to be performed in the assessment programs starting from 2015, suggests the need to evaluate how much the adaptation of the standardized tests to this new medium can provide us with information about the skills of the students in secondary education in a different way to that provided by traditional paper and pencil assessment (PPA). In this study we try to go into this issue in depth through the evaluation of the differences in the results obtained, between PPA and CBA, by Spanish students in maths and reading competence, in order to determine to what extent the two assessment modes reflect the same construct. The results allow us to declare that there is a significant gap in the results of both modes of assessment and that these differences can be explained by factors such as gender, immigrant status, or the socioeconomic and cultural status of the lives of students. Added to this is the evidence regarding variables of effort and access to ICT that lead us to conclude that the relevance of assessment through CBA goes beyond simple differences in the "evaluation mode".

Keywords

Computer-based assessment, paper and pencil assessment, competence, PISA, gap, dispersion.

INTRODUCTION

The development of assessment indicators of the students' academic performance is a key issue to be able to implement improvements in education systems (Battauzetal., 2011, Marcenaro and Vignoles, 2013), since any redesign of educational policies must start from a profound knowledge, *a priori*, of the significance and direction of the relationships between inputs (factors) and products that make up the teaching-learning process. That process is often referred to, from the point of view of economics, as the educational production function.

However, what is the product obtained in this productive process? There isn't an obvious answer to that question, so it seems more appropriate to talk about a multidimensional product. However, as an approximation for estimating the parameters of the educational production functions, we often use the results from some type of student assessment as a one-dimensional measurement to characterize the product of the learning process (see a comprehensive review of this issue in Marcenaro, 2013). Paper and pencil have been the basic elements for filling out these assessment tests in different subjects, either through exams held regularly as part of the subjects program or through standardized tests of the kind developed by the International Programme of student assessment (PISA), among others. But in a paperless world, ie: in a society where digital text is overtaking writing on paper in many areas, very diverse assessments based on increasingly sophisticated technological media are gaining ground. In this context the implementation of the tests conducted under the umbrella of PISA is no exception, and it has happened in the PISA assessment of 2012, one of whose most notable developments was the performance of assessment tests in the traditional format, through written tests, for a restricted sample of teenagers (approximately 40% for the Spanish sample), with IT backup (CBA[1]), in all traditional competences; furthermore, this latter type of assessment is included for the competence of problem solving[2].

Originally the main aim of the introduction of the computer-assisted assessment tests was the inclusion of issues that are difficult to take into consideration in an assessment based on paper and pencil. For example, with regard to the assessment of maths competence, while that animations and simulations are included in the CBA, thus reducing the need for text reading, a more objective measurement of the maths ability of students could be made. Another issue to consider is that the rapid development of computational tools has enhanced the accuracy and speed of performing mathematical calculations, in all their variations, which means these elements have become inseparable travelling companions in professional activity (Hoyles et al., 2002, Cater et al., 2010). So in PISA 2012, after the inclusion of CBA, the importance of maths competences in their application through electronic devices is explicitly acknowledged.

Along with these advantages, assessment through tests using computers may have an additional set of beneficial consequences, among which its lower potential cost in the medium term (Poggio et al., 2005) stands out, since there is no need for printed copies, while at the same time it facilitates assessment through a more flexible questionnaire design, for example through tests which, depending on the successive student responses, can self-select items allowing for a more accurate evaluation of the competences of the

[1] We will use the term "CBA" (Computer Based Assessment) in reference to the assessment tests performed by computer, and the acronym "PPA" (*Paper and Pencil Assessment*) to the assessment tests performed with pencil and paper.
[2] The assessment of this competence was carried out only with CBA, so it was not possible to compare the gaps between PPA and CBA in this competence, and so in this research it was not taken into consideration. However Méndez (2014), in this same volume, makes an in depth study of the determinant factors of the scores in this competence.

individual. Moreover it can assist the collection of a greater quantity of information in less time, as well as more directly facilitating its storage and, consequently, obtaining information useful for redesigning education policies more quickly. This greater dynamism is particularly relevant in the context of international assessment programs with a large number of countries involved.

Despite what has been pointed out not everything is positive, since computer-assisted assessment risks distorting the objective measurement of student competence if the result depends, in an "uncontrolled" way, on the mere ability of the person assessed in using computers, and not on their true competence in terms of literacy in the subject being assessed.

So while the implementation of this type of assessment and its consequences may be analyzed in depth they will serve as a framework for the further and more accurate development of these assessments, since as the OECD reports (OECD, 2013a) in its report "... PISA 2012 represents only the starting point for the possibilities of computer-assisted assessment of maths competence".

As described in Chapter 2, the PISA tradition of CBA implementation started with the 2009 edition, in which 19 of the 65 participating countries (including Spain) were voluntarily included in the international option of ERA (*Electronic Reading Assessment*) which was designed for investigating the performance of students in tasks requiring access, understanding, evaluation and integration of electronic texts in a diverse range of contexts and reading activities (MEC, 2011). This report emphasizes the need to investigate whether digital and printed reading belong to the same construct or not. In other words, it asks whether digital and printed reading measure the same type of skills. This study tries to get to the bottom of this question starting from the analysis of the differences between the results obtained, between PPA and CBA, by Spanish students in maths and reading competence[3], in order to determine to what extent the two assessment modes are picking up on the same type of skills. To do that the characteristics of students and their learning environment will be examined, both at home and at school, in order to analyze whether they may be behind the differences obtained in both types of tests and to what extent some of these characteristics may increase or reduce the gap between them. This objective can be seen to have relevance especially in the context of PISA, in so far as the OECD, who promotes it, seeks to consolidate the CBA as a basic element of their assessments, starting from the next edition of PISA (2015). Also in the event that the differences detected between the two types of assessment are significant, even though they are not large, they could have great relevance especially in contexts in which the results of this type of assessments can be regarded as a reference level (threshold) to pass a subject or competence. This last issue is very well known in Spain, where rates of repetition and dropout are at very high levels[4].

After this brief introduction, the rest of this research study is structured as follows: in the second section there will be a brief review of the previously published relevant literature, in which we will highlight the virtual absence of studies focusing on the description of the differences between assessment methods in Spain. After explaining how

---

[3] In MECD (2013) it is established that the maths competence construct aims "to describe the characteristics of individuals for mathematical reasoning and using concepts, procedures, data and mathematical tools to describe, explain and predict phenomena"; meanwhile reading competence is defined as "understanding, using, reflecting and engaging with written texts, to achieve goals, to develop personal knowledge and potential, and to participate in society."

[4] The early school dropout rate, though it has declined since 2008 by more than 8%, stood at 23.5% in 2013, practically doubling the average of the EU-27 countries (see IVIE, 2013, for a detailed review of this issue).

groundbreaking this research is in relation to the existing contributions in this field of analysis, the data contained in PISA 2012 will be described from a critical perspective in relation to the main stated objective, as a starting point for briefly describing, in the fourth section, the chosen methodology. Starting from this, in the fifth section the estimate of the established models will be addressed, which forms the core of the empirical analyses of this research. Specifically, these analyses will go towards shedding light on the causes that may explain the differences observed in academic performance in the PPA and CBA-type tests, settling how much the two forms of assessment can be considered as measuring the same competence. This paper ends in section six, which presents, in summary, the main conclusions and provides a reflection of the author, based on the obtained results, with some recommendations.

REVIEW OF THE LITERATURE

Without a doubt, the virtual absence of statistical data sources, in the context of the Spanish education system, which allow us evaluate the differences between the PPA and CBA tests, has meant that the relevant literature is very emerging and shallow. So only very recently do we find examples such as the study of Vazquez (2013) which, though focused on data from only ten schools in one Spanish province, shows how "learning processes based on digital reading are still under developed and teachers reveal in most cases a lack of preparation, training and resources available in the school necessary for the effective development of teaching that successfully fosters the competence of digital reading of the students."

In research conducted in the context of other countries with more tradition in the use of Information and Communications Technology (ICT) in assessments of all kinds, as in the case of the United States, the debate on the question of the potential differential effect of performing the same assessment test in both a written and computer-assisted form is becoming increasingly relevant, particularly with regard to the "test mode effect."In this sense the evidence is not only limited but is also far from conclusive. Proof of the low degree of consistency of the results can be found in research such as Alexander et al. (2001) which concludes, based on their empirical evidence, that the only fundamental difference between PPA and CBA tests is that the former were completed in less time (as in Goldberg and Pedula, 2002), but does not find any further difference regarding the potential influence of other characteristics of the students.

Among the first systematic studies devoted to investigating the determinants of differences in PPA and CBA results is that of Gallagher et al. (2000), which finds a slightly negative effect of CBA among girls. However, in the systematic review carried out by Leeson (2006) in relation to the literature on the "assessment mode effect", in which the results obtained in both types of test are classified according to two major families of factors, termed "participatory" (ability, gender or ethnicity) and "technological" (familiarity with the use of computers, type of print in the test, etc.), no systematic patterns are found which allow us to clearly explain the differences in results.

Among the most mentioned results from the previous studies are those of Wallace and Clariana (2002); it can be inferred from that research that students who show a weaker link to the use of computers have worse performance in the written tests and, conversely, those who are classified as stronger students get lower scores in computer-assisted testing. What's more, the model of CBA as a function of the level of skill in the use of computers had a statistically significant correlation even when controlling by performance in PPA, thus emphasizing the influence of familiarity with computers in showing good performance in the CBA tests in the context of maths competence. However in language competence the evidence of Taylor et al. (1999) does not allow us to confirm the existence of a significant correlation between familiarity with computers and the scores obtained in the CBA in a performance assessment test on the knowledge of English as a foreign language(EFL).

No less relevant, regarding the contribution presented here, is the further evidence described by Wallace and Clariana (2002), who allude to the existence of a significant gap between scores in CBA and PPA, in favor of the former, in the group of students with higher academic performance; this difference is not maintained when analyzing the most disadvantaged students, and both Poggio et al. (2005) and Ryan (2013) argue against even the first. Focusing on the comparability of scores in PPA and CBA in relation to maths competence, Bennet et al. (2008), in line with Wallace and Clariana (2002), highlight the lower performance in terms of PPA-type test scores compared with CBA. However in general the existing evidence seems to opt for the superiority of the results of the written tests (PPA) over the tests developed for computer (Sim and Horton, 2005).Nevertheless the evidence from the meta-analysis carried out by Mead and Drasgrow (1993) highlights the high correlation between the two types of assessment.

A different perspective is addressed in Cassady and Gridley(2005) who study the differences between the two modes of assessment not only in terms of the changing habits that may occur among students, but also in relation to the possible psychological repercussions that different assessment formats may lead to. Their main conclusion is that students do not differentiate between the potential results of both types of assessment, and furthermore it states that they feel less "threatened" by an assessment when it is carried out by CBA. The results of Stowell and Bennett (2010) point in the same direction, finding that students who normally experience high levels of anxiety before completing PPA find this decreases when they are examined by CBA.

The empirical studies of Callhoonet al. (2000) or, more recently, Dolan et al. (2005) are equally positive regarding the benefits of the CBA format; in these contributions the usefulness of type CBA assessments is underlined as a tool for promoting a more accurate assessment of students with some type of disability. This question should not be regarded as trivial, despite the fact that the sample of students assessed in PISA excludes disabled students (OECD, 2013b).

It is worth mentioning a final group of contributions which evaluate, in a general way, the potential impact of the availability of ICT in school son the level of student performance in the assessment tests, but without evaluating the differential impact on the assessment modes. In general these studies, among which those of Angrist and Lavy (2002) and Fuchs and Woessman (2004) stand out because of their circulation, find no evidence of a significant causal effect of ICT on student performance, although more recent studies such as those of Machin et al. (2007) and Cabras and Tena (2013) do appear to show some evidence of a positive effect.

STATISTICAL SOURCE

The information contained in PISA 2012, regarding the sample for Spain, makes a reference to the socio-economic, demographic and cognitive characteristics of a total of 25,313 students spread over 902 secondary schools. However, as noted in the introduction, standardized CBA-type tests were only completed by 10,175 teens spread over 368 centers, and these are the focus of this research, since for them comparisons can only be made between the PPA versus CBA assessment modes. Added to this restriction is the fact that the information relating to some potentially highly relevant variables for the teaching-learning process (Marcenaro, 2002) such as, for example, the time spent by students attending classes, and individual work on academic tasks, etc. is only available for two-thirds of the total sample; so the final sample for the empirical analysis carried out consists of 6,546 individuals, a quarter of whom are repeaters.

To better understand the methodology and specifications adopted in the estimates presented in section five it is essential to comment on the result of some of the bivariate descriptive analyses performed at a stage prior to the estimate, which have revealed possible sources for explaining the variations observed in the gap between the written and the computer-assisted tests, both in terms of math sand reading competence. This gap, which will be the variable explained in the estimates, has been calculated for each student as the difference between the "average" plausible value, obtained according to the PISA[5] guidelines, of the results of the paper and pencil test minus the plausible value of the computer-assisted test. Broadly speaking, paying attention to the average value of the variable measuring the PPA-CBA gap, the scores of the assessments conducted with the help of new technologies are lower than those obtained by paper and pencil, except in the case of repeaters in maths competence, where they are the same. However, as can be seen in Figures 4.1 and 4.2, the behavior of this gap, in both competences, shows a decreasing gradient in relation to the scores obtained in CBA, that relationship becoming inverted from, approximately, values close to the CBA average.

In other words, the gap between PPA and CBA gradually reduces as we move towards the "average" student in academic performance, becoming negative (a higher score in CBA than in PPA) for the most gifted students who seem to benefit more from the second form of assessment[6], something which means that in the second part of the estimates we take into account the quartile where students in CBA are found in order to evaluate the gap between this mode of assessment and PPA. This behavior can be observed both among repeater students as well as no repeaters.

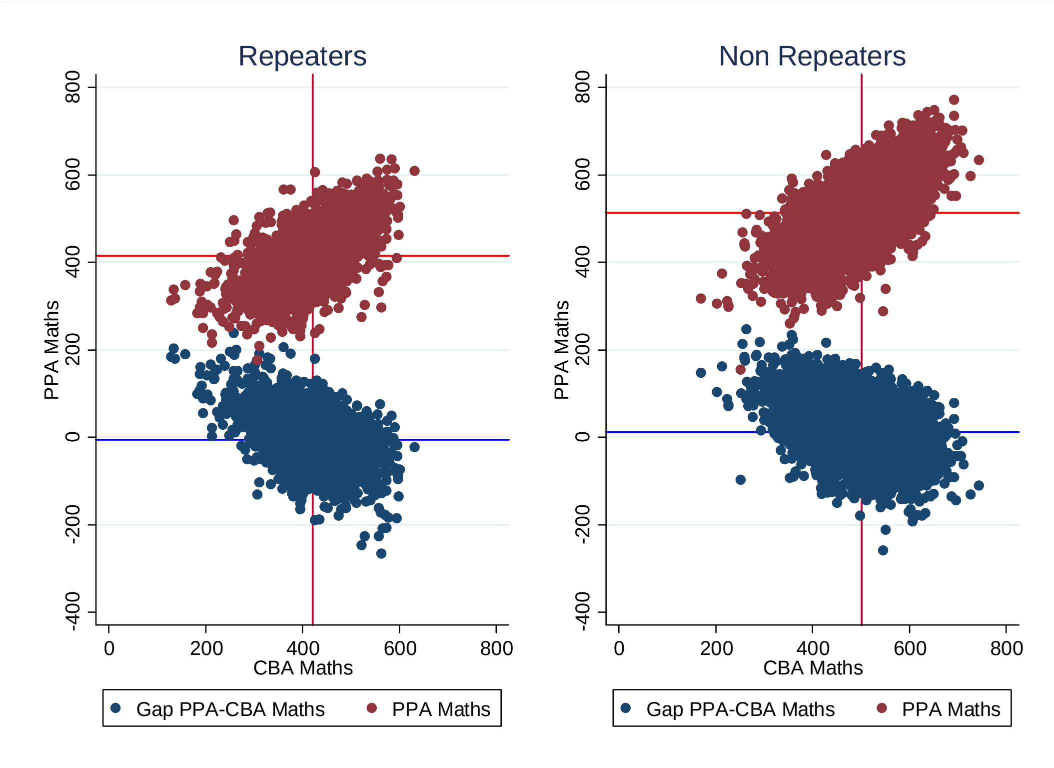
Before continuing it is important to underline that the analyses below have been performed by differentiating according to whether or not the students are repeaters, paying attention, on the one hand, to the results reached by Carabaña (2013) in his analysis of PISA 2012, in which he concludes that "... repetition of a year is associated primarily with cognitive factors" and, on the other hand, to the major differences shown by both groups in scores from tests performed with paper and pencil[7].

Partly following the script from the previous international literature with regard to the potential variables that can affect the PPA-CBA gap in Tables 4.1 and 4.2 (Appendix) we present descriptive statistics of the variables that have attracted the most attention. Thus this gap in the results of both types of test is slightly higher in reading comprehension for women, and in maths for men, mainly because of the relative decline in the results of the girls when CBA is used, which contributes to a reduction of the distance that had been obtained between the sexes, in both competences, when CBA is used. It could be argued

[5] Specifically, given that there are 5 plausible values for each competence assessment, calculations of the statistics are performed for each plausible value and a posteriori we compute the average of the resulting values.
[6] It's worth remembering that Wallace and Clariana (2002) also alluded to the existence of a growing gap for the "best" students.
[7] Furthermore a non-parametric test was carried out on the difference in proportions between the two samples, from which it can be inferred that the proportion of repeater students is significantly lower (25.05%) in the sample that does computer-assisted tests, than in the rest (28.73%); the critical value obtained was 8.89, which leads us to reject the hypothesis of equal proportions at a confidence level of 99%.

that this pattern may be the result of the different usage of computers between boys and girls. However, the contrasts of differences made in the variables relating to the type of computer use (school work at home, leisure, etc.) according to gender do not let us conclude that these are significant.

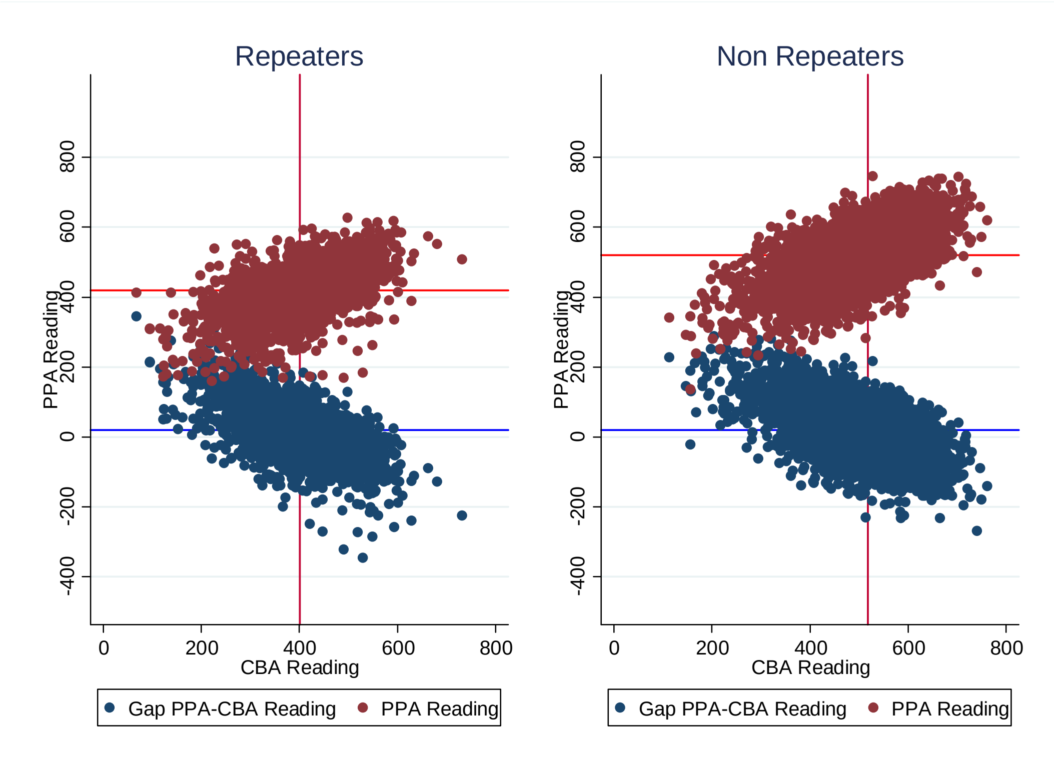
Figure 4.1 Gap between paper and pencil assessment (PPA) and computer-based assessment (CBA) in maths competence



Note: The horizontal lines drawn from the vertical axis represent, respectively, the average scores in PPA Mathematics (red) and the PPA-CBA Mathematics gap (blue).

Source: Authors' own calculation from PISA2012.

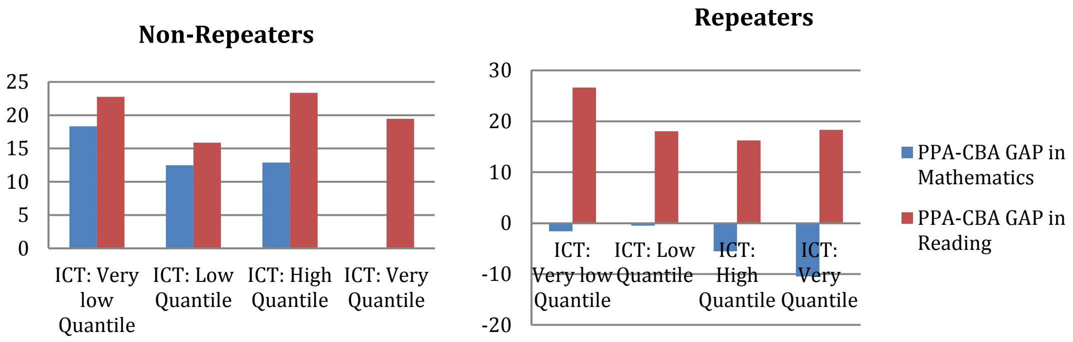
Figure 4.2 Gap between paper and pencil (PPA) and computer-based assessment (CBA) in language competence



Note: The horizontal lines drawn from the vertical axis represent, respectively, the average scores in PPA Language (red) and the PPA-CBA Language gap (blue).
Source: Authors' own calculation from PISA2012.

Familiarity in the use of computers has undoubtedly been one of the most debated issues in this field. However, as argued by McDonald (2002), among others, in so far as computer literacy and availability of electronic devices is becoming universal this factor can be expected to play a decreasingly important role in the results achieved in tests administered by CBA. However, this issue should be examined in depth, so that for the these analyses the information provided by students regarding the availability of technological resources in the school was preferred over that containing the opinion of the school heads, since it was thought that the latter may be more subject to bias. Specifically, in Figure 4.3 we see the average value of the PPA-CBA gap according to the degree of technological development of the school (as expressed by their students through the ICT index). Competence in maths shows a clear trend in the reduction of the gap as the technological resources of the school increase, even turning positive towards the CBA scores when the opinions of repeaters are collected.

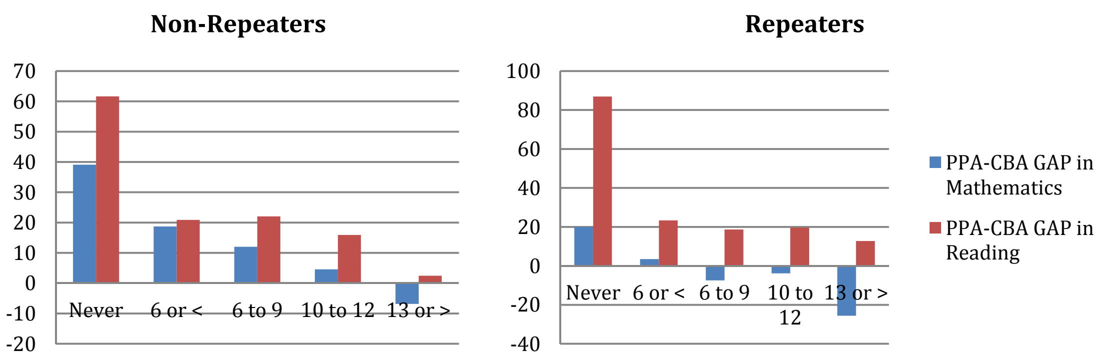
Figure 4.3. Distribution of the gap between PPA-CBA according to the degree of information and communications technology resources revealed by students regarding their school



Source: Authors' own calculation from PISA2012.

In order to complement the analysis of this apparent correlation between ICT development and relatively higher results in CBA than in PPA, we constructed Figure 4.4, whose results should not leave us indifferent, since at first glance they may seem counter-intuitive. Indeed these figures reflect, very clearly in the case of the maths competence, that both among non-repeater as well as repeater students a "late" start in the use of computers is associated with higher scores in CBA, when compared to PPA. At first the opposite may have been expected, assuming that an early start in the use of these technologies is associated with a greater skill in their use. However, from the statistics reported in Table 4.1 (Appendix) it can be inferred that a late start in the use of computers is closely correlated with a low score in PPA and CBA, although the latter drops relatively less compared to PPA, than among the students who start using computers at earlier ages. Consequently, at least from the perspective of the bivariate analysis, the effect on the gap would be better explained by the lesser ability in the standardized PPA tests of the most disadvantaged students than by a greater skill in the use of computers, as scores in CBA also decrease.

Figure 4.4.**Distribution** of the gap between PPA-CBA according to the age at which students began using a computer for the first time



Source: Authors' own calculation from PISA2012.

The performance gap between students that develop in a favorable or advantageous socio-economic and cultural context (high ESCS index) and those who are more disadvantaged is consolidated at positive and high levels, with a distance equivalent to a quarter of a standard deviation, just as speaking a different language at home and at school favors the widening of the gap, in favor of the PPA test, in language competence, as does, to a lesser extent, immigrant status. A similar trend is deduced for the hours of study, whose increase widens the gap, ie: favors the PPA scores over CBA. Alternatively, for the hours of study we use the variable that records the number of complete absences accumulated by the student[8], which go hand in hand, as expected, with a deterioration of the scores in the assessments, in this case higher in PPA than in CBA. Therefore effort, either through more hours of study or through continued attendance seems to favor scores in PPA over CBA.

In order to evaluate to what extent the results collected from the bivariate analysis are maintained when determined simultaneously by the group of variables mentioned in Section 5, we present a set of estimates based on the methodology discussed briefly in the next section.

METHODOLOGY

In the PISA assessment a sample of approximately 35 students per school is used, so some "homogeneity" of the characteristics of students can be expected as a result of belonging to the same school. This is true, a priori, to the extent that attendance at one school or another is determined largely[9] by the distance between the home and the school, so we are likely to observe "common" socioeconomic and cultural patterns among students of the same school, and different ones between schools. This "between" and "within" school double heterogeneity can note accurately captured using the simple techniques of linear regression (OLS), since this methodological approach will provide an inaccurate quantification of the standard errors of the parameters if the students belonging to the same school[10] show similar values in the school variables, grouped in a hierarchical structure, so that the average correlation between variables of students in the same school will be higher than that existing between students from different schools (Hox, 1998); this will lead us to underestimate the standard deviations of the parameters of the estimated regression hyperplane and, therefore, to consider the correlation between the explained variable and any of explanatory variables as significant when sometimes it shouldn't.

So, as an alternative multilevel analysis is normally used (Raudenbush and Bryk, 2002), in which the contribution to the variance (total heterogeneity) of the characteristics of the students is estimated, as a first level, and those added according to school, such as higher hierarchical level, as a second level[11]. This implies the estimation of a regression hyperplane for each school (level 2) and not a single hyperplane for all schools. In algebraic terms, to understand the model that will be finally estimated, we start with the equation (1):

$$Y_{ij} = \alpha_j + X'_{ij}\beta_j + \varepsilon_{ij} \qquad (1)$$

where the matrix Y_{ij} denotes scores in standardized tests in a particular competence for a student "i" who attends school "j"; α_j is a parameter of level that reflects the differential

[8] Data for this variable are available for the entire sample, not just for two-thirds as occurred for the hours of study, so they are used as an alternative in the specifications estimated in Section 5.
[9] This influence may not be as direct in the case of private schools.
[10] The same reasoning applies when information is available regarding the grouping in classrooms-classes within a school. Unfortunately there is not such detailed information in PISA.
[11] The number of observations required, per level, to put it into practice with guaranteed robustness in the estimates is controversial. However Maas and Hox (2005) state that, in the case where we want to accurately calculate the variance decomposition within and between schools, we need at least 100 observations from the first level and 10 from the second; this criterion is met in these estimates.

effect on Y_{ij} of each school, in other words, a proxy of the "quality" of the school, X_{ij} is a vector of student characteristics, and ε_{ij} is the term of idiosyncratic error.

Following Angrist and Pischke (2010), if we have an uneven number of observations per school, with a low number of young students sampled in some schools and a relatively high number in others, it is appropriate to impose some structure to undertake the estimate, for example by using random effects (via Feasible Generalized Least Squares), or to use the method implemented by Cabras and Tena (2013), which involves the application of a nonparametric Bayesian estimation model[12].

Equation (1) assumes that the intercept of the regression is the same for all transverse units. However, it is very likely that we need to control the "individual" character of each state. The random effects model suggests that each school has a different level parameter (interception term), ie: it is not fixed but behaves as a random variable with an average value (α) suffering deviations by an amount represented by the random variable u_i. So that:

$$\alpha_j = \alpha + u_j \qquad (2)$$

Substituting in (1) we get:

$$Y_{ij} = \alpha + \beta_1 X_{1ij} + u_j + \varepsilon_{ij} \qquad (3)$$

This is the expression of the random effects model. In this case the hypothesis of zero correlation between the disturbance term associated with the school and the exogenous variables (X) could not be rejected so the random effects estimations preferable[13]. If we go a step further, and add the effect on the scores (Y_{ij}) of the variables considered at the school level (Z_j) to the expression (3) we come to the following model:

$$Y_{ij} = \alpha + \beta X_{ij} + \delta Z_j u_i + \varepsilon_{ij} \qquad (4)$$

which represents the multilevel model of estimation using random effects. This methodology has become widespread in the analysis of PISA data, especially in the reports accompanying the latest editions (see for example: García-Montalvo, 2013, or Mediavilla and Escardibul, 2014).

Moreover, as a consequence of the profile shown in Figures 4.1 and 4.2, it is necessary to complement the analysis based on the previous methodology with some specification that determines the value of the PPA-CBA gap according to the relative situation of the students in the distribution of scores in CBA. This could be done directly by including the score in CBA in the expression (3) as an additional regressor, either continuously or by using quartiles[14], but this strategy would lead us to have to solve a serious problem of endogeneity of the latter variable. Alternatively, as has been done, it could be the estimation of the multilevel model for each of the sub-samples that come from fragmenting the distribution of scores in CBA into quartiles. This strategy is a first approximation to the issue, and so it will be addressed in greater depth in future research.

---

12 In this study and in the references contained therein a detailed analysis is presented of the difficulties in distinguishing between correlation and causality relationships in the context of educational production functions.
13 The advantage of using random effects was determined by applying the Hausman test.
14 To take into account possible non-linearities in the potential association between the gap and CBA.

RESULTS

Starting from the methodological strategy that has been described we proceeded to estimate the coefficients of the multilevel analysis[15] for the complete sample of students who completed both PPA as well as CBA; the results are provided in Tables 4.3 and 4.4 (Appendix). In these we show two specifications, in the second of which the hours of attendance has been substituted by the number of whole days that the student missed class, in order to analyze the consistency of the results of the first specification when the problem of the sample reduction is avoided, which involves the consideration of the variables of study time as a regressor (available only for two thirds of the sample).

The values and degree of statistical significance of the coefficients are such that the PPA-CBA gap is reduced at a higher rate for women than for men when assessing the behavior in maths competence of the non-repeaters, from which it could be deduced that the computer medium, when it is determined in the analysis by many other characteristics of the students, favors a convergence in results between boys and girls, since on average in PPA the latter did better than their female classmates (by 23 standardized points)[16]. Interestingly, despite the fact that among repeater students boys also show better outcomes, the CBA-PPA gap is not significantly different as a consequence of another factor that will be discussed later. In order to take into consideration the potential restriction that is the reference quartile in CBA maths where each teenager is found, Tables 4.5 and 4.6 (Appendix) were generated, for competence in maths and language, respectively, of the non-repeaters[17], which complements the previous analysis in the sense that the significant reduction in the gender gap is stable regardless of the "ability" shown by the scores in CBA. When we move on to Table 4.4 we see that for language competence the gap is not reduced but increased in all cases between girls and boys, and that this increase is most striking among the female students who have the worst performance in the CBA (Table 4.6). It could be interpreted that conducting tests in CBA format "favors" teenage girls over their male classmates, helping to improve their relative position in the competence for which they are most disadvantaged in PPA, and leaving the advantage they already had in language competence unchanged or even slightly increased, especially among those who are "less gifted".

The negative effect for immigrant status (grouping those of first and second generation) that is typically observed in the literature with respect to the results in the standardized PPA-type tests (Calero and Escardibul, 2013) is maintained when the PPA-CBA gap in language competence is analyzed, so that it may be concluded that it contributes even more to a reduction, compared to the native students, of the scores in PPA more than in CBA. The opposite occurs in maths, a competence in which the gap is accentuated between repeater students as a result, at least partially, of the greater proportion of immigrant students among the repeaters.

On the other hand, the potential correlation between immigration and diglossia is not relevant[18], since the latter refers to the use of a different language at home and at school, in line with that observed by Carabaña (2013), when maths is assessed in PPA, and also exhibits a statistically insignificant coefficient when the correlation is established with

---

[15] The models were estimated using *Stata* software, version 13.1.

[16] See Marcenaro and López (2013) for a review of the recent literature on the differences in educational performance according to gender.

[17] The respective tables for repeater students have not been included for reasons of space and in order to focus our attention on students whose chances of transition to higher levels of education are higher.

[18] Although the proportion of students who speak a different language at home and at school is much higher among the immigrant students (45.05%) than among native students (15.21%).

the CBA-PPA gap. In contrast for the language competence the relative contribution of diglossia means that the relative reduction in this gap due to immigrant status is reversed for those who speak different languages inside and outside home, but only among students with a lower performance in CBA (quartiles I and II, see Table 4.6). Considering the decomposition in quartiles the overall effect of the two variables is only positive, and sizeable, in maths (approximately 18 points). This differential behavior according to skills and quartiles, in the case of language, requires an in-depth reflection because the results of the implementation of CBA type assessments can bring up important nuances in the imbalances seen between Autonomic Regions in terms of assessments, if attention is paid to their different starting points in terms of diglossia.

A key variable in the analyses of the educational production function is the socioeconomic and cultural status (ESCS) in the lives of the students, as this explains a substantial percentage of the variance observed in PPA in the various competences assessed in PISA (see this volume in Cordero et al., 2014, for a detailed analysis). At first they addressed the estimations presented in the tables of the Appendix using different alternative proxies to this status, such as the highest level of education reached by the father and the mother, number of books at home, or the presence of specific books of humanities, but none of them showed substantially different effects when included in alternative specifications, being highly correlated with the ESCS index (which PISA provides as a derived variable); so the latter was used to make the estimations more parsimonious.

Despite the relevance of the effect of the ESCS index on the PPA scores, when this is included in the model that tries to explain the PPA-CBA gaps in the form of quartiles in order to collect possible non-linearities, its influence is weak in maths competence, balancing itself out among the repeater students. This contribution of the higher status of students from the highest quartiles of ESCS increases the PPA-CBA gap in language competence; therefore, although they get higher PPA scores, the substitution of this mode of assessment for that of CBA should reduce the distance with respect to students found in "more disadvantaged" environments in terms of ESCS, the results turning out to be robust when determined by the level of performance in CBA, where they are classified (Table 4.6).

As for the hours of study, like other variables mentioned below (guided work hours, help time provided by parents, etc.), the question was answered by two thirds of the sample, since the PISA 2012 questionnaire included 28 questions, to be completed by all of the sampled students, and three differentiated sets of additional questions, from which students were randomly assigned to answer two[19]. Among the students who were asked to answer these questions (specification I of the different tables of estimations) a positive gap of around 6 points is estimated between those who spend 3 or more hours of study per week, both in maths as well as language competence, compared to those who spend fewer hours, and therefore make less effort towards academic work outside school. This distance disappears between repeaters in language competence but is much more marked in maths. One possible explanation is that being hours of study which are not specifically spent with the help of technological support contributes more to the traditional mode of learning,

[19] We compared whether the subset of students who responded to the question about hours of study, from among those who were selected for having been assessed either using PPA as well as CBA, showed any differences in their individual characteristics compared to those who were not asked to answer the question from among those selected, and no statistically significant difference was found, something which guarantees the randomness of the subsample.

linked to PPA, than that which is assess able by CBA. This suggests that the skills captured in PPA and CBA are not exactly the same.

With the aim of going into further detail about the above question an alternative specification was estimated in which study time was replaced by the number of hours, using discreet variables for hours, which the teenager claims to devote each week to study using a computer[20]; their coefficients were not significant in language, but in the maths competence the gap is reduced by 30 points (equivalent to half a standard deviation) if we compare those who spend 7 or more hours with those who do not spend any. Furthermore, when the impact of this last variable on the CBA scores is estimated we get coefficients which are negative and increasing with the number of hours[21], so we can conclude that there is a selection bias in the use of computers for studying among lower performance students, since they get lower scores in both types of assessments, and this effect reduces their PPA results more than their CBA results. To sum up, as far as the main objective of this research is concerned, the help of computers for study does not make students more competent in general, but using this mode of examination could mean lower performance students "artificially" cut back the disadvantage that they show in PPP-type tests, without implying a better performance in the teaching-learning process.

Regarding this "non-independent" learning, ie: guided academic work, performed outside the school, the number of hours devoted to this negatively affects standardized scores in CBA, as well as its influence on the maths gap (it is negligible in reading) which is reduced[22]. Similarly the time spent by students performing academic tasks with a personal tutor outside school also shows an increasing negative trend with the number of hours in the CBA assessment of both competences. Although in this latter case the narrowing of the PPA-CBA gap is not significant. Finally, as to the academic work done at home, time spent with the help of a parent or other family member doing school work at home also shows a negative and increasing correlation with the PPA[23] and CBA scores, as well as an insignificant effect on the gap between both of them. In Marcenaro (2013) it is argued that the potential adverse influence on PPA could be explained, on the one hand, by the limited confidence of parents in the pupils given the previous results of their children, leading them to try to help them in their academic tasks outside school, or on the other hand, by a "passivity" which would arise in the students receiving such help, which would contribute to poorer results in these standardized tests. This line of argument could also apply to behavior regarding CBA, which would result in the lack of effect on the gap. From what has been said it appears that, in general, the "supervised" learning outside the school has enough impact to tip the balance in favor of one or the other mode of assessment.

Another group of variables included in the estimations reflects the level of ICT development of the school and the student. Among the available variables, the preliminary statistical analysis carried out (see Section 2) hinted that the availability and/or use of a tablet device in the school maintained a correlation with PPA-CBA gap. Its inclusion in the multivariate conditional-multilevel- analysis leads us to affirm that students who claim to have a tablet available for their use in school get a significantly lower score in PPA than those who do not, and this difference is more substantial when they use it actively and

[20] The corresponding tables with the coefficients of these estimates are not shown for reasons of space, but they are available upon request to the author.

[21] Another argument, which seems less plausible, is that students with lower average productivity, measured by the relationship between effort in terms of time commitment and results, use the computer to try to compensate for their slow learning, which would explain the concentration of students who use the computer for academic tasks among the least advantaged in terms of scores.

[22] The corresponding tables with the coefficients of these estimationss are not shown for reasons of space but are available upon request to the author.

[23] Marcenaro (2013) shows comparable evidence when the exam results of students with administrative data are evaluated.

towards maths comprehension[24]. Thus the reduction in the gap of the scores between written and computer-assisted tests is highly determined by the negative impact of this availability on performance in the written tests, in line with what is observed for the time spent using a computer for doing schoolwork. It is also interesting to note that this effect is diluted if the availability of the device refers to the student's home.

The other variable to consider, within this set, is that concerning the starting age of the use of computers; the estimated coefficient simply that the earlier they begin to use computers, the greater the gap in favor of PPA in maths competence; in language competence the correlation is not obvious. It should be emphasized that the positive relationship with the gap in maths is very high (one third of a standard deviation) in the highest quartiles of the distribution of scores in CBA (Table 4.5), therefore it is among the most gifted students that the contribution becomes more tangible in both the CBA as well as with greater emphasis in PPA. This relationship corroborates what was discussed in section 2, where we speculated on the potential reasons that may explain this result.

A final set of variables that have been considered to explain the potential gaps between PPA and CBA could be grouped under the label of "self-esteem" or confidence in the abilities of each student[25]. Therefore they are considered as variables that can be classified as proxy of the students' own perception of their ability. Firstly five dummy variables are included in the estimations, constructed from the response on a Likert scale with five values (from "Strongly Agree" to "Strongly disagree"), which reflects whether the assessed students think they give up quickly when faced with problems. If they answer that they disagree their score is increased both in PPA as well as in CBA, but the first to a greater degree, resulting in a widening of the gap, especially in the maths competence. Although the variable reflecting "whether difficulty discourages you" shows the same trend of correlation with PPA and CBA, yet leaves the gap unchanged. The same applies in relation to the variables: I maintain interest in something I start (perseverance), do not finish until everything is perfect (perfectionism) and make more effort than expected.

The second part of the estimations presented in Tables 4.3 to 4.6 of the Appendix shows the effect of second-level variables, ie: variables whose effect is added at a school level, in order to differentiate within this effect the part due to the heterogeneity of the students, within each school, which can be attributed to that of the set of schools. In relation to this the results from research on previous editions of PISA show that there is a relatively small contribution of schools to the variation of the scores of the students (about 20%).of a relatively small contribution of schools (about 20%), to a variation of the scores of students. In the research presented here and resulting from the estimates we have obtained Charts 4.1 to 4.4 (Appendix). In their first rows these charts include the coefficients of intra-class correlation (ICC), which are calculated from the other two elements shown in the charts below: "Standard deviation of the constant across schools" (the greater it is, the greater the difference between schools, so it is more appropriate to use random effects, as these will be greater) and "Standard deviation of the term of total error"[26]. The ICC value depends largely on the variability of the observed values: the more

[24] It might be thought that the presence of tablets in the school might respond to the type of school, classified by the total amount of public funding received, but the descriptive statistics showed that this is not so, since the proportion of students who stated that they have tablet devices at their schools are very similar in private, semi-private and public schools, although it is slightly higher in the latter.

[25] The estimated coefficients for the variables included in the corresponding alternative specifications are not shown for reasons of space.

[26] Specifically, the intra-class coefficient of correlation results from the application of the following formula: Variance of the constant through schools / (Variance of the constant through schools + Variance of the term of total error).

homogeneous the sample studied, the lower the value of the ICC tends to be. We can see that in the case of the gaps in PPA-CBA scores, both in maths as well as reading and repeaters as well as non-repeaters, they show an intra-class variation which is not at all negligible and significant in all cases, indicating that differences between schools affect the gaps presented by students in the maths and reading tests. The same can be applied to the ICC estimations that divide the sample into quartiles[27].

However the greater weight of these variations remains with the individual characteristics of students. So if we look at the added variables at a school level few variables stand out as significant in explaining the PPA-CBA gap, particularly in the case of maths comprehension where only the growth of the student-teacher ratio means that the gap increases, and that only by a small amount. In language competence the student/teacher ratio is also relevant, and even more the proportion of immigrants in the school and the rate of use of ICT in the school. The latter widens the gap substantially among non-repeaters.

Conclusions

The use of ICT in the teaching-learning process can, in principle, contribute to an improvement in performance in computer based tests, facilitating the adaptation of content, but it can also be detrimental to learning if it contributes to students becoming distracted and losing focus on content. Whichever the effect, what is really relevant, in terms of the aim of this research, is whether this technology widens or narrows the potential gap between the scores performed in standardized tests by students when, for filling them out, paper and pencil is substituted by the screen and mouse.

One of the most controversial points in the limited existing literature (Puhan et al., 2007), in relation to the gap between the results in PPA and CBA is that which sets out the possibility that what the scores in this latter "assessment mode" really reflect is the students' ability to use technological media rather than the competence itself which is being assessed (maths, etc.). The descriptive statistics provided showed feeble evidence since the gap was reduced and even reversed, placing the results in CBA above those of PPA, for students with less relative digital literacy (if this is approximated from the age when the use of computers starts). However, in other indicators, such as the perception of students on the ICT resources of the school, it showed the opposite behavior. Despite the apparent lack of consistency of the results from the bivariate analysis, the multivariate conditional regression analysis has shed a lot of light on the matter.

In particular it could be interpreted that conducting the tests in a CBA format "favors" the female students over their male classmates, in helping to improve their relative position in the maths competence, in which they have historically shown a lower average performance than boys, measured in terms of PPA, and leaving the advantage they already had in language competence unchanged or even slightly increased.

On the other hand the combination of effects of the variable which classifies the students as immigrants together with that of diglossia requires us to extend our consideration of the benefit of the tests in CBA mode to the field of the differences between Autonomic Regions, due to significant imbalances between them in terms of the relevance of diglossia and the shifting composition of their populations. If we extrapolate

[27] These results therefore support the use of multilevel regressions.

this reasoning to the other countries making up the international PISA sample the argument can gain even more weight.

No less important is the fact that the substitution of the pencil and paper assessment by the CBA would reduce the distance with respect to the students found in "more disadvantaged" environments (in terms of ESCS), which leads back to the hypothesis that differences between PPA and CBA go beyond the simple assessment mode effect. But these are not the only results that bring us closer to this idea, since the greater relative contribution of study hours (as a measure of the effort of the students) to the PPA scores over CBA, and the smaller gap estimated for students who spend more hours studying with the help of a computer, imply the existence of a selection bias among the "less gifted", which are those who get the greatest relative performance in CBA.

In summary, at least two things can be inferred from the results reported in this investigation. Namely, on the one hand, the help of ICT in the teaching-learning process, inside or outside the school, does not make students more competent in general, but using an exam mode based on these technologies could mean that the "less gifted" students "artificially" reduce the disadvantage they show when performing PPA-type tests, without implying a better performance in the teaching-learning process. Secondly, and as a consequence of the above, if, as has been assumed up to now, assessment using pencil and paper properly approximated the level of student learning, these new technologies, in changing the course of the students' scores in CBA compared to PPA, would be measuring different skills which are more present among those who interact more with these new technologies.

Finally, to underline that this research is only a starting point, since the research on this issue has a lot of future relevance to something as complex as the measurement of the results of the teaching and learning processes, which are the foundations for the creation of knowledge and, consequently, the social, cultural and economic development of a region.

References

ALEXANDER, M.W., BARTLETT, J. E., TRUELL, A. D. Y OUWENGA, K. (2001) "Testing in a computer technology course: An investigating of equivalency in performance between online and paper and pencil methods". *Journal of Career and Technical Education*, 18, 69-80.

ANGRIST, J. Y LAVY, V. (2002) "New evidence on classroom computers and pupil learning". *The Economic Journal*, 112, 735-765.

BATTAUZ, M. R., BELLIO, R. Y GORI, E. (2011) "Covariate error adjustment for multilevel models with application to educational data". *Journal of Educational and Behavioral Statistics*, 36(3), 283-306.

BENNETT, R. (2003) "Online Assessment and the Comparability of Score Meaning, Educational Testing Service, Princeton, New Jersey".www.ets.org/Media/Research/pdf/RM-03-05-Bennett.pdf.

BENNETT, R. E., BRASWELL, J., ORANJE, A., SANDENE, B., KAPLAN, B. Y YAN, F. (2008) "Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP". *Journal of Technology, Learning, and Assessment*, 6(9).

CABRAS, S. Y TENA, J. D. (2013) "Estimación del efecto causal del uso de ordenadores en los resultados de los estudiantes en el test PISA 2012". En: INEE (Ed.), PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Análisis secundario.

CALHOON, M. B., FUCHS, L. S. Y HAMLETT, C. L. (2000) "Effects of computer-based test accommodations on mathematics performance assessments for secondary students with learning disabilities". *Learning Disability Quarterly*, 23, 271-282.

CALERO, J. Y ESCARDIBUL, O. (2013) "El rendimiento del alumnado de origen inmigrante en pisa 2012". En: INEE (Ed.), PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Análisis secundario.

CARABAÑA, J. (2013) "Repetición de curso y puntuaciones pisa ¿cuál causa cuál?." En: INEE (Ed.), PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen I: Resultados y contexto (páginas del capítulo citado).

CASSADY, J. C. Y GRIDLEY, B. E. (2005) "The effects of online formative and summative assessment on test anxiety and performance". *Journal of Technology, Learning, and Assessment*, 4(1).

CATER, K., ROSE, D., THILLE, C. Y SHAFFER, D. (2010) "Innovations in the classroom". Presentation at the Council of Chief State School Officers (CCSSO) National Conference on Student Assessment, Detroit MI.

CORDERO, J., PEDRAJA, F. Y SIMANCAS, R. (2014) "Superando las barreras: factores determinantes del rendimiento en escuelas y estudiantes con un entorno desfavorable". En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

DOLAN, R. P., HALL, T. F., BANERJEE, M., CHUN, E. Y STRANGMAN, N. (2005) "Applying principles of universal design to test delivery: The effect of computer-based read-aloud on test performance of high school students with learning disabilities". *Journal of Technology, Learning, and Assessment*, 3.

FUCHS, T. Y WOESSMAN, L. (2004) "Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at Schools". *Brussels Economic Review*, 47, 359-389.

GALLAGHER, A., BRIDGEMAN, B. Y CAHALAN, C. (2000) "The effect of computer-based tests on racial/ethnic, gender, and language groups". GRE Board Professional Report No. 96–21P. Princeton, NJ: Education Testing Service.

GOLDBERG, A. Y PEDULLA, J. J. (2002) "Performance differences according to test mode and computer familiarity on a practice GRE". *Educational and PsychologicalMeasurement*, 62(6), 1053-1067.

GARCÍA-MONTALVO, J. (2013) "Crisis, igualdad de oportunidades y resultados educativos en España: una visión retrospectiva desde PISA 2012". En: INEE (Ed.), PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen I: Resultados y contexto (páginas del capítulo citado).

HOX, J. J. (1998) "Multilevel modeling: When and why". En: I. Balderjahn, R. Mathar, Y M. Schader (Ed.), Classification, data analysis and data highways (pp. 147-154). New York: Springer.

HOYLES, C., WOLF, A., MOLYNEUX-HODGSON, S. Y KENT, P. (2002) "Mathematical skills in the workplace: final report to the Science Technology and Mathematics Council", Project Report, Institute of Education, University of London, Science, Technology and Mathematics Council, London, http://eprints.ioe.ac.uk/1565/1/Hoyles2002Mathematical Skills.pdf.

IVIE (2013) "El abandono educativo temprano: análisis del caso español". Instituto Valenciano de Investigaciones Económicas, Valencia.

LEESON, H. V. (2006) "The Mode Effect: A Literature Review of Human and Technological Issues in Computerized Testing". *International Journal of Testing*, 6 (1), 1-21.

MAAS, C. J. M. Y HOX, J. J. (2005) "Sufficient sample sizes for multilevel modelling". *Methodology*, 1, 86-92.

MACHIN, S., S. MCNALLY Y O. SILVA (2007) "New Technology in Schools: Is There a Payoff?," *Economic Journal*, 117(522), 1145-1167.

MASON, B., M. PATRY Y D. BERSTEIN (2001) "An examination of the equivalence between non-adaptive computer based and traditional testing". *Journal of Education Computing Research*, 24(1), 29-39.

MARCENARO, O. (2002) "Los estudiantes universitarios y el uso de su tiempo: una función de producción educativa". Tesis Doctoral, Universidad de Málaga.

MARCENARO, O. (2013) "El rendimiento del alumnado andaluz a examen". Editado por la Consejería de Presidencia e Igualdad de la Junta de Andalucía, Sevilla.

MARCENARO, O. Y LÓPEZ, L. A. (2013). "El efecto del capital cultural sobre el rendimiento educativo diferencial por género". Actas del IV Congreso de Economía Feminista, Carmona.

MARCENARO, O. Y VIGNOLES, A. (2013). "The quality measures of educational performance: an empirical approach". Actas del XVI Encuentro de Economía Aplicada, Granada.

MCDONALD, A. S. (2002) "The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments". *Computers & Education*, 39, 299–312.

MEAD, A. D. Y DRASGOW, F. (1993) "Equivalence of computerized and paper-and-pencil cognitive ability tests: a meta analysis". *Psychological Bulletin*, 114, 449-458.

MEDIAVILLA, M. Y ESCARDIBUL, J.O. (2014) "Efecto de las TICs en la adquisición de competencias. Un análisis de género y titularidad de centro para las evaluaciones por ordenador." En: INEE (Ed.), PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Resolución de problemas (páginas del capítulo citado). Madrid: Autor.

MÉNDEZ, I. (2014) "Factores determinantes del rendimiento en resolución de problemas. España en perspectiva internacional". En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MINISTERIO DE EDUCACIÓN (2011) "PISA - ERA 2009. Programa para la Evaluación, Internacional de los Alumnos, OCDE. Informe español". Instituto de Evaluación, Ministerio de Educación. Madrid.

MECD (2013) "Marcos y pruebas de evaluación de PISA 2012". Madrid.

OECD (2013a) "PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy". Paris.

OECD (2013b) "PISA 2012 Results: What Makes a School Successful (Volume IV) Resources, Policies and Practices: Resources, Policies and Practices". Paris.

POGGIO, J., GLASNAPP, D. R., YANG, X. Y POGGIO, A. J. (2005) "A comparative evaluation of score results from computerized and paper and pencil mathematics testing in a large scale state assessment program". *Journal of Technology, Learning, and Assessment*, 3(6).

PUHAN, G., BOUGHTON, K. Y KIM S. (2007) "Examining differences in examinee performance in paper and pencil and computerized testing". *Journal of Technology, Learning, and Assessment*, 6(3).

RAUDENBUSH, S. W. Y BRYK, A. S. (2002) "Hierarchical linear models: Applications and data analysis methods". 2nd edition. Newbury Park , CA, Sage.

RYAN, B. (2013) "Computer-Based Versus Paper-Pencil Modes of Administration United States Government End of Course Exams: Student Cumulative Grade Point Averages as Predictors of Success". Ph dissertation. Lindenwood University.

SIM, G. Y HORTON, M. (2005) "Performance and attitude of children in computer based versus paper based testing". Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA) (pp. 3610-3614). Chesapeake, VA: AACE.

STOWELL, J. R. Y BENNETT, D. (2010) "Effects of online testing on student exam performance and test anxiety". *Educational Computing Research*, 42(2), 161-171.

TAYLOR, C., KIRSCH, I., EIGNOR, D. Y JAMIESON, J. (1999) "Examining the relationship between computer familiarity and performance on computer-based language tasks". *Language Learning*, 49, 219–274.

UN (2013) World Telecommunication/ICT Indicators database 2013 (17th Edition). http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.

VÁZQUEZ, E. (2013) "Análisis y evaluación de la didáctica con base en lectura digital en el marco PISA: el caso de España". *Revista Iberoamericana de Evaluación Educativa*, 6(1), 61-76.

WALLACE, R. Y CLARIANA, P. (2002) "Paper-based versus computer-based assessment: key factors associated with the test mode effect". *British Journal of Educational Technology*, 33(5), 593-602.

APPENDIX

Table 4.1Descriptive statistics for the characteristics of the non-repeater students

	PPA Maths		PPA Reading		CBA Maths		CBA Reading		PPA-CBA Maths		PPA-CBA Reading	
Variables	Average	St.Dev.	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.
Female	500.7	2.5	527	2.3	489.5	3.5	505.1	3.4	11.2	3.4	22	3.8
Male	526.4	2.9	507.8	3.2	513.4	3.5	489.8	5	13	3.5	18	4.9
Immigrant	468.8	6.2	480.5	6.7	448.4	6.4	449.4	8.3	20.4	6.1	31.1	7.7
Native	516.6	2.3	520.8	2.4	505.2	3.1	501.3	3.9	11.4	3.3	19.5	4.2
Diglossia: Yes	498.4	3.5	507.6	4.8	487.5	7.2	471.8	9.7	10.9	6.9	35.8	10.8
Diglossia: No	518	2.3	521.3	2.3	505.3	2.8	506.1	3.8	12.7	2.8	15.2	3.5
ESCS (Very low quartile)	484	3	493.5	3.2	476	4.3	474.4	5.5	8	4.2	19.1	5
ESCS (low quartile)	507.8	2.6	512.7	2.7	498.1	3.9	487.9	5.5	9.7	3.3	24.7	5.6
ESCS (High quartile)	520.5	4.2	524.1	4	508.2	3.5	503.9	4.3	12.3	3.8	20.2	4.5
ESCS (Very High quartile)	547.4	2.9	547.1	4.2	527.9	4.3	530.9	4.6	19.5	4.3	16.2	5.3
Hours of study (0 to 2 hours)	498.4	3.8	496.9	3.4	489.4	4.2	478.7	5.2	9	3.5	18.2	5
Hours of study (3 to 5 hours)	511.8	3.5	514.4	3.1	498.7	4.1	492.5	6.2	13.1	4.9	21.9	6.3
Hours of study (6 to 8 hours)	519	3.3	526.6	3.8	502.2	4.1	502	4.6	16.9	3.6	24.6	4.9
Hours of study (9 to 30 hours)	525.1	2.7	539.9	2.9	515.6	3.5	516.6	4	9.4	3.5	23.3	4.6
Miss a whole class: Never	519.8	2.2	522.3	2.3	508	3.2	501.3	4	11.8	3.3	21.1	4.2
Miss a whole class: 1 to 2 times	492.2	3.3	504.5	3.7	476.6	4.3	485.4	4.6	15.6	3.7	19.1	4.9
Miss a whole class: 3 to 4 times	493.1	8.6	499.7	10.2	511.2	12.8	510.5	11.2	-18.1	13.8	-10.8	10.6
Miss a whole class: 5 or more times	454.4	15.5	455.4	20.7	452	12.4	448.2	19.3	2.5	13	7.2	9.9
Single parent home: Yes	507.8	4.7	514.2	4.6	506.2	5.9	493.6	6.5	1.5	4.7	20.6	5.5
Single parent home: No	514.3	2.2	517.9	2.5	501.3	3.2	497.2	3.9	13	3.3	20.7	4.2
Tablet in the school: Yes and they use it	475.1	11.6	462.5	11.5	478	11.9	440.4	20.5	-2.9	5.8	22.1	12.8
Tablet in the school: Yes but they don't use it	495.1	7.6	493.6	7.8	506.1	9.7	486.2	14	-11	9	7.4	13.5
Tablet in the school: No	516	2.3	521.9	2.4	501.9	3.2	501.3	3.9	14	3.1	20.6	4.1

Age when started to use the computer: Never	508.6	29.3	471.7	28.8	469.5	45.1	410.1	48.3	39.1	32.1	61.6	50.5
Age when started to use the computer: 6 or<	525.2	2.9	528.5	3.5	506.5	3	507.6	4.5	18.7	3.4	20.9	5
Age when started to use the computer: 6 - 9	515	2.8	518.7	2.7	502.9	3.9	496.7	4.5	12	3.4	22	4.4
Age when started to use the computer: 10 - 12	495.1	3.3	507.5	3.6	490.5	4.3	491.6	5	4.6	4.4	15.9	4.5
Age when started to use the computer: 13 or>	458.9	9.5	470.6	7.5	465.8	8	468.1	9.1	-6.8	6.8	2.4	7.2
ICT Quartile: Very low	512.8	3.9	514.8	3.8	494.5	4.5	492	5.5	18.3	4.4	22.8	5.4
ICT Quartile: Low	519.4	2.4	526.4	2.7	507	3.7	510.5	4.3	12.5	3.6	15.9	4
ICT Quartile: High	517.9	3.5	525.3	3	505	4.3	501.9	4.9	12.9	3.5	23.4	5
ICT Quartile: Very high	492.4	4.4	494.8	5.6	492.3	5.4	475.3	8.6	0.1	5	19.5	7.5
Private school	541	5	544.3	6.3	520.4	8.9	514.7	9.2	20.6	8.1	29.6	9.3
Semi-private school	525.6	4.7	526.7	5	511.5	5.9	513.3	7.5	14.1	6.2	13.4	9.3
Public school	503.3	3	509.9	3.1	493.3	4.2	487.9	6	10	3.8	22	5.3

Source: From PISA 2012 (the sample only includes students who performed the PPA and CBA tests.

Table 4.2 Descriptive statistics for the characteristics of the repeater students

	PPA Maths		PPA Reading		CPA Maths		CPA Reading		PPA-CBA Maths		PPA-CBA Reading	
	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.
Female	402.2	2.7	433.6	3	412.8	4.1	414.3	6.3	-10.6	3.6	19.4	6.8
Male	424.7	2.9	409.8	3.8	426.9	3.4	390.1	5.5	-2.2	3.4	19.7	4.4
Immigrant	404.4	7.9	409.9	6.1	397.4	5	386.6	7.4	7	7.2	23.4	10
Native	417.2	2.1	422.2	3	426.7	3	402.4	5.6	-9.5	3.1	19.8	4.7
Diglossia: Yes	413.5	5.4	418.8	5.4	418.8	4.9	384	7.6	-5.3	5.8	34.8	8
Diglossia: No	415	2.4	420.5	3	420.9	3.2	405.3	6.3	-5.9	3.2	15.2	5.7
ESCS (Very low quartile)	399.2	3.3	408.1	3.9	407.9	4.1	387.5	5.9	-8.7	3.9	20.6	6.4
ESCS (low quartile)	408.1	4.4	413.8	4.8	412.4	4.2	394.3	8.1	-4.3	4.8	19.5	8.4
ESCS (High quartile)	424.8	3.9	431.4	5.2	429	4.8	409.6	7.3	-4.1	3.5	21.7	6.2
ESCS (Very High quartile)	430.6	4.9	431	4.5	436.5	4.4	414.2	7.7	-5.8	4.7	16.8	5.9
Hours of study (0 to 2 hours)	395.2	3.5	395.6	4	410.5	4.6	392.1	6.5	-15.3	3.9	3.5	6.2
Hours of study (3 to 5 hours)	419.2	10.5	428.5	8.7	414.1	6.5	396.8	14.3	5.1	10.9	31.7	13.4
Hours of study (6 to 8 hours)	431.5	3.9	431.1	4.1	434.6	4.1	415.5	6.9	-3.1	4.6	15.5	5.7
Hours of study (9 to 30 hours)	432.5	3.8	445.4	5	434.3	5.8	423.2	8.5	-1.8	5.9	22.2	7.5
Miss a whole class: Never	423.8	2.8	424.9	3	426.2	3.8	398.7	5.4	-2.4	3.4	26.2	5
Miss a whole class: 1 to 2 times	407.5	4.4	419	5	414.2	2.9	409.2	7.7	-6.7	3.7	9.8	7.9
Miss a whole class: 3 to 4 times	386.5	9.4	403.4	9.4	416.2	11.1	400.4	11.1	-29.7	11.7	3	9.2
Miss a whole class: 5 or more times	379.4	9.6	387.8	11	398.6	15.7	346.6	13.3	-19.2	11	41.2	12.9
Singleparent home: Yes	422.5	4.2	431.6	4.8	430.7	5.1	416.2	9.4	-8.3	4.7	15.4	10.9
Singleparent home: No	414.8	2.5	418.9	3.2	419.8	2.9	396.7	5.3	-5	3.3	22.2	5
Tablet in the school: Yes and they use it	385.1	6.7	374.8	9.1	415.6	7.6	369.9	12.6	-30.5	4.3	4.9	8.2
Tablet in the school: Yes but they don't use it	391.5	12.9	397.5	11.7	398.6	11.6	373.1	17	-7.1	11.8	24.4	14.9
Tablet in the school: No	420.9	2.1	428.3	2.9	422.6	2.7	406.7	5.5	-1.7	3.1	21.6	5.3
Age when started to use the computer: Never	338	28.8	294.5	27.5	318	23.3	207.6	56.9	20	23.1	86.9	30.6
Age when started to use the computer: 6 or <	427.3	3.9	427.5	5	423.9	4.2	404.2	8.4	3.5	3.3	23.3	8.7
Age when started to use the computer: 6 - 9	419.3	3.4	427.3	3.5	426.7	3.8	408.6	6.4	-7.5	4.2	18.7	5.6
Age when started to use the computer: 10 - 12	403.1	4.7	414.6	4.1	407	4.8	395	6.3	-3.9	4.8	19.6	6.2

Age when started to use the computer: 13 or >	390.6	11.4	393.6	10.8	416.2	12.5	380.9	17.7	-25.6	9.4	12.8	12.1
ICT Quartile: Very low	413.7	3.4	422.9	4.6	415.3	5.1	396.3	7.8	-1.6	3.8	26.6	8.2
ICT Quartile: Low	426.9	4.6	434.7	5.1	427.4	3.7	416.7	6.8	-0.5	3.1	18	6.6
ICT Quartile: High	420.4	4.1	425.6	4.9	426	5.8	409.3	8.6	-5.5	6.1	16.2	6.4
ICT Quartile: Very high	402.7	4.9	402.8	6.3	413.1	4.9	384.5	6.7	-10.5	5.1	18.3	6.9
Private school	448	9.4	452.4	9	433.8	9.9	421	6.8	14.1	7.9	31.3	8.1
Semi-private school	437.7	5.1	435.1	6.5	439.7	7.3	421.3	10.8	-2	6.1	13.8	9
Public school	408.4	2.8	415.6	3.2	416.1	3.1	395.4	6.8	-7.7	3.8	20.1	6.5

Source: From PISA 2012 (the sample only includes students who performed the PPA and CBA tests.

Table 4.3 Multilevel estimates to explain the gap between PPA and CBA in maths competence

Variables	Non Repeaters		Repeaters	
	spec. I	spec. II	spec. I	spec. II
FEMALE	-5.857***	-3.343***	-1.744	-1.910
	(1.356)	(1.044)	(3.107)	(2.353)
IMMIGRANT	14.231***	15.201***	23.848***	23.180***
	(3.171)	(2.503)	(4.873)	(3.634)
DIGLOSSIA	0.382	-0.632	-1.853	-0.163
	(2.248)	(1.780)	(5.237)	(3.884)
ESCS. Reference category: Very low				
ESCS (Low quartile)	-1.117	-2.238	-3.113	0.332
	(2.087)	(1.632)	(3.816)	(2.856)
ESCS (High Quartile)	-0.317	0.353	-7.283	-5.283
	(2.063)	(1.640)	(4.428)	(3.400)
ESCS (Very high quartile)	3.245	4.618***	-6.215	-4.413
	(2.161)	(1.719)	(5.862)	(4.531)
HOURS OF STUDY. Reference category: (0 to 2 hours) / MISS A WHOLE CLASS Reference category: (Never missed a whole class)				
Hours of study (3 to 5 hours) / Missed the whole class (1 to 2 times)	5.875***	2.476	16.707***	6.677**
	(1.840)	(1.526)	(3.750)	(2.674)
Hours of study (6 to 8 hours) / Missed the whole class (3 to 4 times)	6.311***	-27.896***	16.068***	-7.644
	(1.998)	(4.664)	(4.896)	(5.601)
Hours of study (9 to 30 hours) / Missed the whole class (5 or more times)	5.978***	-25.019***	13.252***	-9.670
	(2.045)	(6.905)	(4.583)	(7.114)
SINGLE PARENT HOME	-5.960**	-5.597***	-1.116	-3.462
	(2.403)	(1.866)	(4.407)	(3.311)
TABLET IN THE SCHOOL. Reference category: (No tablet at school)				
There is a tablet and they use it	-13.953***	-18.002***	-14.922***	-17.860***
	(4.235)	(3.179)	(5.792)	(4.592)
There is a tablet but they don't use it	-7.223*	-8.522***	-18.019**	-12.072**
	(3.977)	(3.091)	(7.525)	(5.633)
AGE AT WHICH STARTED TO USE A COMPUTER. Reference category (13 or older)				
Six years or less	9.810*	13.877***	-10.468	5.497
	(5.380)	(4.146)	(7.682)	(5.923)
Six to nine years	8.463	11.742***	-13.074*	3.532
	(5.327)	(4.107)	(7.498)	(5.753)
Ten to twelve years	0.718	4.293	-13.738*	1.254
	(5.473)	(4.217)	(7.746)	(5.880)
TYPE OF SCHOOL. Reference category: (Public school)				
Private school	5.003	2.320	4.140	14.594
	(16.084)	(15.816)	(21.254)	(19.030)
Semi-private school	4.435	-1.815	0.753	10.503
	(14.221)	(14.002)	(18.697)	(16.587)
PERCENTAGE OF GIRLS IN THE SCHOOL	-32.547	-32.821	-26.556	-34.845
	(23.455)	(23.060)	(35.223)	(30.324)
AVERAGE STUDIES OF THE FATHER	3.505	4.924*	7.825**	6.753**
	(2.636)	(2.561)	(3.365)	(2.904)
AVERAGE STUDIES OF THE MOTHER	0.137	-1.600	-1.554	-1.174
	(2.572)	(2.503)	(3.209)	(2.887)
PROPORTION OF IMMIGRANTS	0.106	-0.038	-0.197	-0.244

	(0.195)	(0.186)	(0.218)	(0.199)
RATIO STUDENTS/TEACHER	2.062***	1.848***	0.897	1.238
	(0.704)	(0.691)	(1.046)	(0.868)
FLAG RATIO STUDENTS / TEACHER	6.317	4.349	2.408	4.536
	(14.843)	(14.597)	(21.415)	(18.312)
USE OF ICT IN THE SCHOOL	4.077	6.091	4.798	9.650
	(5.394)	(5.272)	(6.818)	(6.180)
USE OF ICT IN MATHS LESSONS	1.143	0.461	-0.786	-3.434
	(5.374)	(5.271)	(6.927)	(6.303)
GROUPING FOR THE CLASSES. Reference category (No grouping)				
Grouping in all classes	4.953	8.013	-0.529	4.307
	(8.029)	(7.896)	(10.610)	(9.622)
Grouping in some classes	4.953	8.013	-0.529	4.307
	(8.029)	(7.896)	(10.610)	(9.622)
ADMISSION BECAUSE PARENTS SHARE PHILOSOPHY OF THE SCHOOL. Reference category: (This criterion isn't used)				
This criterion is always used	-4.169	-4.057	2.305	-0.085
	(6.233)	(6.118)	(8.469)	(7.630)
This criterion is sometimes used	7.279	5.394	4.940	0.814
	(6.891)	(6.759)	(9.155)	(8.191)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO APPOINT/ FIRE TEACHERS	-16.873 (13.427)	-12.626 (13.223)	-4.400 (17.396)	-16.663 (15.434)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET SALARIES AND SALARY INCREASES	-1.817 (7.064)	-0.923 (6.948)	-3.931 (10.369)	-5.530 (9.184)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO DECIDE BUDGETS AND BUDGET ALLOCATION	8.282 (11.355)	11.468 (11.116)	20.739 (14.438)	21.630 (13.235)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET ASSESSMENT RULES, DISCIPLINE AND STUDENT SELECTION	29.752 (21.601)	28.149 (21.338)	14.518 (28.454)	29.801 (27.452)
Constant	-76.314**	-54.778	-69.284	-55.202
	(37.638)	(34.870)	(50.598)	(44.841)
Observations	4,338	6,591	1,170	1,796
Number of schools	343	343	308	319
Wald test	169.47***	284.77***	124.52***	152.65***

*** denotes variable significant to level 1%; ** to 5%; * to 10%.

Source: from PISA 2012microdata.

(1) 14 dummy variables are included corresponding to the autonomic regions with an extended sample and a dummy which corresponds to the students resident in the three autonomic regions without an extended sample.

Table 4.4 Multilevel estimates to explain the gap between PPA and CBA in language competence.

Variables	Non Repeaters		Repeaters	
	spec. I	spec. II	spec. I	spec. II
FEMALE	2.312	3.680***	3.749	4.895*
	(1.513)	(1.187)	(3.616)	(2.815)
IMMIGRANT	-9.191***	-9.549***	-4.379	-5.367
	(3.539)	(2.845)	(5.666)	(4.348)
DIGLOSSIA	18.184***	19.758***	16.421***	18.059***
	(2.513)	(2.026)	(6.120)	(4.654)
ESCS. Reference category: Very low				
ESCS (Low quartile)	3.029	3.839**	4.438	3.735
	(2.330)	(1.855)	(4.453)	(3.419)
ESCS (High Quartile)	6.827***	8.222***	3.547	1.231
	(2.302)	(1.864)	(5.158)	(4.069)
ESCS (Very high quartile)	6.333***	6.856***	1.275	1.858
	(2.411)	(1.954)	(6.845)	(5.426)
HOURS OF STUDY. Reference category: (0 to 2 hours) / MISS A WHOLE CLASS Reference category: (Never missed a whole class)				
Hours of study (3 to 5 hours) / Missed the whole class (1 to 2 times)	5.141**	-12.422***	0.157	-13.453***
	(2.053)	(1.735)	(4.371)	(3.200)
Hours of study (6 to 8 hours) / Missed the whole class (3 to 4 times)	9.862***	-15.911***	5.259	-7.147
	(2.229)	(5.301)	(5.711)	(6.702)
Hours of study (9 to 30 hours) / Missed the whole class (5 or more times)	8.564***	-12.048	6.734	9.840
	(2.284)	(7.848)	(5.334)	(8.511)
SINGLE PARENT HOME	-2.419	-0.210	7.974	5.728
	(2.681)	(2.121)	(5.133)	(3.962)
TABLET IN THE SCHOOL. Reference category: (No tablet at school)				
There is a tablet and they use it	-5.417	-6.563*	-9.090	-10.514*
	(4.730)	(3.615)	(6.733)	(5.497)
There is a tablet but they don't use it	-7.465*	-6.792*	-17.256**	-2.581
	(4.439)	(3.514)	(8.779)	(6.740)
AGE AT WHICH STARTED TO USE A COMPUTER. Reference category (13 or older)				
Six years or less	-0.976	-8.063*	18.251**	21.387***
	(6.005)	(4.713)	(8.951)	(7.083)
Six to nine years	6.325	-2.133	17.446**	18.852***
	(5.946)	(4.669)	(8.741)	(6.879)
Ten to twelve years	5.878	-3.158	12.353	17.860**
	(6.108)	(4.794)	(9.025)	(7.030)
TYPE OF SCHOOL. Reference category: (Public school)				
Private school	11.152	13.447	-14.425	-0.759
	(20.089)	(19.992)	(27.522)	(23.703)
Semi-private school	-2.161	-2.436	-21.164	-12.712
	(17.766)	(17.698)	(24.165)	(20.660)
PERCENTAGE OF GIRLS IN THE SCHOOL	-16.310	-14.544	11.659	22.155
	(29.315)	(29.165)	(45.745)	(37.694)
AVERAGE STUDIES OF THE FATHER	0.365	0.643	9.817**	6.694*
	(3.274)	(3.224)	(4.373)	(3.612)

AVERAGE STUDIES OF THE MOTHER	0.018 (3.195)	-0.139 (3.151)	-7.720* (4.175)	-6.311* (3.597)
PROPORTION OF IMMIGRANTS	0.584** (0.240)	0.685*** (0.233)	0.654** (0.285)	0.578** (0.248)
RATIO STUDENTS/TEACHER	1.676* (0.879)	1.619* (0.874)	3.222** (1.356)	2.843*** (1.078)
FLAG RATIO STUDENTS / TEACHER	16.240 (18.547)	12.679 (18.459)	55.415** (27.818)	48.537** (22.776)
USE OF ICT IN THE SCHOOL	11.984* (6.723)	12.364* (6.655)	11.913 (8.859)	8.089 (7.699)
USE OF ICT IN MATHS LESSONS	6.645 (6.706)	5.360 (6.660)	6.046 (9.020)	8.312 (7.857)
GROUPING FOR THE CLASSES. Reference category (No grouping)				
Grouping in all classes	-1.026 (10.036)	-1.160 (9.988)	-15.070 (13.733)	-11.197 (11.977)
Grouping in some classes	2.704 (6.924)	3.817 (6.862)	9.280 (9.068)	7.416 (7.927)
ADMISSION BECAUSE PARENTS SHARE PHILOSOPHY OF THE SCHOOL. Reference category: (This criterion isn't used)				
This criterion is always used	-0.837 (7.785)	1.242 (7.736)	-1.182 (10.979)	-1.610 (9.501)
This criterion is sometimes used	-3.241 (8.600)	-4.667 (8.539)	-4.085 (11.890)	-5.389 (10.207)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO APPOINT/ FIRE TEACHERS	4.001 (16.773)	-0.975 (16.712)	13.926 (22.488)	3.366 (19.229)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET SALARIES AND SALARY INCREASES	-10.104 (8.832)	-9.757 (8.790)	-8.960 (13.379)	-5.003 (11.413)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO DECIDE BUDGETS AND BUDGET ALLOCATION	-5.854 (14.172)	-6.095 (14.049)	-11.959 (18.784)	-17.041 (16.507)
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET ASSESSMENT RULES, DISCIPLINE AND STUDENT SELECTION	27.602 (27.025)	33.014 (26.993)	67.643* (38.113)	55.464 (34.516)
Constant	-30.665 (46.042)	9.558 (43.649)	-66.021 (64.692)	-30.422 (55.592)
Observations				
Number of schools	343	343	308	319
Wald test	165.72***	261.51***	92.35***	116.67***

*** denotes variable significant to level 1%; ** to 5%; * to 10%.

Source: from PISA 2012 microdata.

(1) 14 dummy variables are included corresponding to the autonomic regions with an extended sample and a dummy which corresponds to the students resident in the three autonomic regions without an extended sample.

.

Table 4.5 Multilevel estimates to explain the gap between PPA and CBA, for the different quartiles of results of students in the CBA of maths competence (sample of non-repeaters)

Variables	Quartile I		Quartile II		Quartile III		Quartile IV	
	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II
FEMALE	-7.090**	-6.164***	-10.670***	-7.040***	-10.961***	-6.484***	-9.250***	-5.373**
	(2.981)	(2.325)	(3.003)	(2.216)	(2.648)	(2.029)	(2.756)	(2.124)
IMMIGRANT	12.269**	12.443***	✓	10.234*	✓	✓	27.476**	✓
	(5.095)	(3.944)		(5.427)			(11.759)	
DIGLOSSIA	✓	✓	✓	✓	✓	✓	✓	✓
ESCS. Reference category: Very low								
ESCS (Low quartile)	✓	✓	✓	✓	8.106*	✓	✓	✓
					(4.240)			
ESCS (High Quartile)	✓	✓	✓	7.028**	6.861*	✓	✓	✓
				(3.373)	(4.158)			
ESCS (Very high quartile)	✓	✓	✓	8.913**	16.412***	11.355***	✓	7.057*
				(3.525)	(4.374)	(3.546)		(3.765)
HOURS OF STUDY. Reference category: (0 to 2 hours) / MISS A WHOLE CLASS Reference category: (Never missed a whole class)								
Hours of study (3 to 5 hours) / Missed the whole class (1 to 2 times)	6.793*	✓	14.388***	-6.935**	✓	✓	✓	✓
	(3.832)		(3.971)	(3.216)				
Hours of study (6 to 8 hours) / Missed the whole class (3 to 4 times)	8.469**	-17.137*	15.480***	-33.568***	✓	✓	✓	-55.791***
	(4.215)	(9.567)	(4.385)	(10.567)				(10.231)
Hours of study (9 to 30 hours) / Missed the whole class (5 or more times)	8.503*	✓	18.116***	✓	6.749*	-34.819**	✓	✓
	(4.409)		(4.563)		(3.904)	(13.636)		
SINGLE PARENT HOME	✓	✓	✓	✓	-8.150*	✓	-8.116*	-7.761**
					(4.826)		(4.584)	(3.557)
TABLET IN THE SCHOOL. Reference category: (No tablet at school)								
There is a tablet and they use it	-16.527**	-18.977***	✓	-21.052***	-20.893**	-26.283***	-21.955**	-15.673*
	(7.243)	(5.407)		(6.083)	(10.445)	(7.721)	(10.644)	(8.096)
There is a tablet but they don't use it	-13.332*	-11.870**	✓	✓	✓	-10.630*	-25.431***	-22.230***

	(7.952)	(5.925)				(5.914)	(9.676)	(7.460)
AGE AT WHICH STARTED TO USE A COMPUTER. Reference category (13 or older)								
Six years or less	✓	15.837**	22.182*	19.690**	✓	✓	35.091*	25.769**
		(6.826)	(11.887)	(8.499)			(19.772)	(11.767)
Six to nine years	✓	12.584*	20.314*	16.467*	✓	✓	33.569*	23.476**
		(6.738)	(11.825)	(8.447)			(19.702)	(11.718)
Ten to twelve years	✓	✓	✓	✓	✓	✓	✓	✓
TYPE OF SCHOOL. Reference category: (Public school)								
Privates chool	✓	✓	✓	✓	✓	✓	✓	✓
Semi-private school	✓	✓	✓	✓	✓	✓	✓	✓
PERCENTAGE OF GIRLS IN THE SCHOOL	-58.772*	✓	✓	✓	✓	-33.877*	✓	✓
	(34.086)					(20.181)		
AVERAGE STUDIES OF THE FATHER	✓	6.384**	✓	✓	4.599*	4.369*	5.608*	7.231***
		(3.200)			(2.647)	(2.412)	(3.001)	(2.659)
AVERAGE STUDIES OF THE MOTHER	✓	✓	✓	✓	✓	✓	✓	✓
PROPORTION OF IMMIGRANTS	✓	✓	✓	-0.343*	✓	✓	✓	✓
				(0.200)				
RATIO STUDENTS/TEACHER	1.552*	✓	1.716**	1.244*	1.512**	1.467**	1.310*	1.215*
	(0.860)		(0.675)	(0.641)	(0.700)	(0.643)	(0.784)	(0.721)
FLAG RATIO STUDENTS / TEACHER	✓	✓	✓	✓	✓	✓	✓	✓
USE OF ICT IN THE SCHOOL	✓	✓	✓	✓	✓	✓	12.126**	9.149*
							(6.089)	(5.440)
USE OF ICT IN MATHS LESSONS	✓	✓	✓	✓	✓	✓	✓	✓
GROUPING FOR THE CLASSES. Reference category (No grouping)								
Grouping in all classes	✓	✓	✓	✓	✓	✓	17.765*	16.043*
							(9.489)	(8.479)
Grouping in some classes	✓	✓	✓	✓	-11.942**	✓	✓	✓
					(5.390)			
ADMISSION BECAUSE PARENTS SHARE PHILOSOPHY OF THE SCHOOL								
Reference category: (This criterion isn't used)								
This criterion is always used	✓	✓	✓	✓	✓	✓	✓	✓
This criterion is sometimes used	✓	✓	15.616**	12.159*	✓	✓	✓	✓

			(6.983)	(6.513)				
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO APPOINT/ FIRE TEACHERS	✓	✓	✓	✓	✓	✓	✓	✓
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET SALARIES AND SALARY INCREASES	✓	✓	✓	✓	✓	✓	✓	✓
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO DECIDE BUDGETS AND BUDGET ALLOCATION	30.343**	26.011**	✓	✓	✓	✓	✓	✓
	(13.939)	(13.197)						
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET ASSESSMENT RULES, DISCIPLINE AND STUDENT SELECTION	✓	✓	✓	✓	✓	✓	✓	✓
Constant	-44.045	-42.587	-215.39***	-91.619**	-116.491**	-73.159	-68.621	-39.410
	(49.882)	(43.507)	(56.988)	(43.634)	(52.449)	(50.067)	(42.410)	(43.421)
Observations	1,085	1,650	1,086	1,649	1,083	1,650	1,084	1,642
Number of schools	289	307	317	327	296	312	267	291
Wald Test	100.59***	138.67***	146.78***	158.97***	126.87***	125.01***	122.75***	189.56***

*** denotes variable significant to level 1%; ** to 5%; * to 10%.

Source: from PISA 2012 microdata.

(1) 14 dummy variables are included corresponding to the autonomic regions with an extended sample and a dummy which corresponds to the students resident in the three autonomic regions without an extended sample.

Table 4.6 Multilevel estimates to explain the gap between PPA and CBA, for the different quartiles of results of students in the CBA of language competence (sample of non-repeaters)

Variables	Quartile I		Quartile II		Quartile III		Quartile IV	
	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II
FEMALE	6.901**	7.967***	5.637*	9.436***	✓	8.439***	✓	4.627**
	(3.257)	(2.460)	(2.926)	(2.292)		(2.225)		(2.281)
IMMIGRANT	-12.633**	✓	-20.981***	-15.620***	✓	-16.815***	✓	-21.263***
	(6.406)		(6.807)	(5.481)		(6.006)		(8.133)
DIGLOSSIA	10.328**	12.145***	12.428**	12.181***	✓	✓	✓	8.378*
	(4.684)	(3.529)	(4.856)	(3.907)				(4.450)
ESCS. Reference category: Very low								
ESCS (Lowq uartile)	✓	✓	10.948**	9.128***	✓	✓	9.906*	12.691***
			(4.436)	(3.471)			(5.260)	(4.206)
ESCS (High Quartile)	✓	6.503*	17.912***	16.264***	✓	✓	18.554***	19.473***
		(3.487)	(4.402)	(3.547)			(5.037)	(4.038)
ESCS (Very high quartile)	✓	8.286**	17.642***	17.877***	✓	10.218***	22.366***	23.394***
		(4.028)	(4.715)	(3.774)		(3.655)	(5.066)	(4.065)
HOURS OF STUDY. Reference category: (0 to 2 hours) / MISS A WHOLE CLASS Reference category: (Never missed a whole class)								
Hours of study (3 to 5 hours) / Missed the whole class (1 to 2 times)	7.405*	-11.046***	13.478***	-13.550***	✓	-10.441***	✓	-13.220***
	(4.020)	(3.261)	(4.008)	(3.284)		(3.444)		(3.496)
Hours of study (6 to 8 hours) / Missed the whole class (3 to 4 times)	13.425***	✓	20.905***	✓	12.815***	✓	13.093***	-26.982***
	(4.513)		(4.299)		(4.278)		(4.623)	(10.444)
Hours of study (9 to 30 hours) / Missed the whole class (5 or more times)	10.691**	-27.745**	21.925***	✓	20.214***	-34.829**	13.810***	✓
	(4.766)	(12.382)	(4.467)		(4.328)	(14.775)	(4.749)	
SINGLE PARENT HOME	✓	✓	✓	✓	✓	✓	✓	✓
TABLET IN THE SCHOOL. Reference category: (No tablet at school)								
There is a tablet and they use it	✓	-14.526***	-23.576**	-22.273***	✓	✓	✓	✓
		(5.093)	(9.426)	(7.524)				
There is a tablet but they don't use it	✓	✓	✓	✓	-20.989**	-14.353*	-18.349**	-13.970*

					(9.919)	(7.775)	(9.225)	(7.365)
AGE AT WHICH STARTED TO USE A COMPUTER. Reference category (13 or older)								
Six years or less	✓	✓	✓	✓	25.282* (13.543)	✓	✓	✓
Six to nine years	✓	✓	✓	✓	30.300** (13.462)	✓	✓	✓
Ten to twelve years	✓	✓	✓	✓	27.568** (13.769)	✓	✓	✓
TYPE OF SCHOOL. Reference category: (Public school)								
Private school	✓	✓	✓	✓	✓	✓	✓	✓
Semi-private school	✓	✓	✓	✓	✓	✓	✓	✓
PERCENTAGE OF GIRLS IN THE SCHOOL	✓	✓	✓	✓	✓	✓	✓	✓
AVERAGE STUDIES OF THE FATHER	✓	✓	✓	✓	✓	✓	✓	✓
AVERAGE STUDIES OF THE MOTHER	✓	✓	✓	✓	✓	✓	✓	✓
PROPORTION OF IMMIGRANTS	✓	0.510** (0.247)	✓	✓	✓	✓	✓	✓
RATIO STUDENTS/TEACHER	1.774* (1.046)	✓	1.706** (0.791)	1.593** (0.728)	2.034*** (0.695)	1.618** (0.664)	✓	1.366* (0.718)
FLAG RATIO STUDENTS / TEACHER	✓	✓	✓	✓	34.684** (14.739)	✓	✓	✓
USE OF ICT IN THE SCHOOL	✓	✓	✓	✓	✓	✓	11.346* (6.476)	12.902** (5.756)
USE OF ICT IN MATHS LESSONS	✓	✓	✓	✓	✓	✓	✓	✓
GROUPING FOR THE CLASSES. Reference category (No grouping)								
Grouping in all classes	✓	✓	✓	✓	✓	✓	✓	✓
Grouping in some classes	✓	✓	✓	✓	✓	✓	✓	✓
ADMISSION BECAUSE PARENTS SHARE PHILOSOPHY OF THE SCHOOL. Reference category: (This criterion isn't used)								
This criterion is always used	✓	✓	✓	✓	✓	✓	✓	✓
This criterion is sometimes used	✓	✓	✓	✓	✓	✓	✓	✓

INDEPENDENCE OF HEAD/TEACHERS/BOARD TO APPOINT/ FIRE TEACHERS	✓	✓	✓	✓	✓	✓	✓	✓
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET SALARIES AND SALARY INCREASES	✓	✓	✓	✓	✓	✓	✓	✓
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO DECIDE BUDGETS AND BUDGET ALLOCATION	✓	✓	✓	✓	✓	✓	✓	✓
INDEPENDENCE OF HEAD/TEACHERS/BOARD TO SET ASSESSMENT RULES, DISCIPLINE AND STUDENT SELECTION	✓	60.118* (34.613)	✓	✓	✓	✓	✓	✓
Constant	-40.961 (58.383)	-8.909 (49.867)	-45.137 (39.508)	53.724 (57.566)	-196.24*** (56.573)	-71.809 (44.701)	-30.167 (47.763)	-39.658 (41.356)
Observations	1,085	1,649	1,084	1,647	1,087	1,652	1,082	1,643
Number of schools	275	293	315	327	299	316	267	286
Wald Test	77.40***	110.93***	141.86***	150.29***	130.18***	126.11***	108.58***	150.51***

*** denotes variable significant to level 1%; ** to 5%; * to 10%.

Source: from PISA 2012 microdata.

(1) 14 dummy variables are included corresponding to the autonomic regions with an extended sample and a dummy which corresponds to the students resident in the three autonomic regions without an extended sample.

Chart4.1 Random values of the multilevel regression models for the complete sample: gap in Maths

	Non- Repeaters		Repeaters	
	spec. I	spec. II	spec. I	spec. II
Coefficient of intra-class correlation	0.409	0.419	0.338	0.360
	0.024	0.022	0.035	0.029
Standard deviation of the constant through schools	34.458	34.564	33.615	34.167
	1.627	1.557	2.351	1.976
Standard deviation of the term of total error	41.461	40.730	47.083	45.529
	0.466	0.365	1.129	0.836

Chart 4.2 Random values of the multilevel regression models for the complete sample: gap in Reading

	Non- Repeaters		Repeaters	
	spec. I	spec. II	spec. I	spec. II
Coefficient of intra-class correlation	0.470	0.475	0.426	0.388
	0.023	0.022	0.035	0.029
Standard deviation of the constant through schools	43.567	44.020	46.718	43.324
	1.965	1.919	2.978	2.454
Standard deviation of the term of total error	46.220	46.280	54.213	54.370
	0.518	0.415	1.313	1.000

Chart 4.3 Random values of the complete multilevel regression models by quartiles: gap in Maths (subsample of the non-repeater students)	Quartile I		Quartile II		Quartile III		Quartile IV	
	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II
Coefficient of intra-class correlation	0.396	0.388	0.255	0.303	0.274	0.287	0.345	0.321
	0.038	0.033	0.039	0.031	0.037	0.031	0.041	0.033
Standard deviation of the constant through schools	34.277	33.347	24.760	26.745	23.329	24.077	27.726	26.494
	2.447	2.127	2.262	1.816	1.969	1.672	2.278	1.882
Standard deviation of the term of total error	42.297	41.862	42.362	40.599	37.993	37.979	38.239	38.512
	1.066	0.813	1.097	0.798	0.962	0.739	0.963	0.746

Chart 4.4 Random values of the complete multilevel regression models by quartiles: gap in Language (subsample of the non-repeater students) Variables	Quartile I		Quartile II		Quartile III		Quartile IV	
	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II	spec. I	spec. II
Coefficient of intra-class correlation	0.427	0.423	0.370	0.375	0.259	0.292	0.336	0.319
	0.035	0.030	0.037	0.031	0.038	0.031	0.040	0.034
Standard deviation of the constant through schools	40.139	38.567	31.983	32.511	24.682	26.669	29.891	28.763
	2.590	2.194	2.213	1.974	2.179	1.847	2.414	2.068
Standard deviation of the term of total error	46.508	45.018	41.725	41.999	41.763	41.512	42.062	42.033
	1.143	0.858	1.069	0.825	1.060	0.809	1.050	0.812

5. Effect of ICT in the acquisition of competences. An analysis of gender and types of schools for assessments by computer[1]

Mauro Mediavilla

Universitat de València & IEB & GIPE

Josep-Oriol Escardíbul

Universitat de Barcelona & IEB & GIPE

SUMMARY

This chapter analyzes the effects that the possession and use of information and communication technologies (ICT) have on the acquisition of competences assessed by computer in PISA 2012. The analysis is developed in a segmented way according to gender (and partly, according to type of school). The main results are as follows. Firstly, the ICT variables have a greater effect on the maths assessment than in the rest of the competences assessed. Secondly, some ICT factors, analyzed individually, show a divergent behavior in some cases. In general, the positive impact shown by variables associated with the existence of ICT (at home or at school), such as in maths, does not correspond with the impact of the use or time spent using these (which has a negative sign in most competences, maybe due to the presence of negative causality). Thirdly, the variables that show students' relationship with ICT are significant. Thus, the enjoyment of using computers (as

[1] We thank the Ministry of Education, Culture and Sport for funding this research. We also thank the cooperation received from all members of INEE and, in particular, Ismael Sanz Labrador, Ruth Martin Escanilla, Francisco García Crespo and Luis Sanz San Miguel. All errors or omissions are the responsibility of the authors.

entertainment) is positively related to the results in computer-based tests, as well as an early age at which students start using ICT. Fourthly, results show a greater effect of the ICT variables in computer-based tests than the results indicated by the studies that analyze the effect of this type of variables in the traditional on-paper assessments. Finally, the segmented analysis by gender and type of school proves to be relevant.

Keywords

Computer-Based Assessment, Gender, PISA 2012, ICT, Types of school

INTRODUCTION

The PISA (*Programme for International Student Assessment*) evaluation, developed by the OECD, analyzes 15 year old students' competences in four areas: reading literacy, maths, sciences and financial knowledge (the latter for the first time). In each edition of the assessment one of the aforementioned competences is assessed in depth, corresponding to maths in 2012. Regarding the application of digital technology in the test, Spain has been a pioneer by participating in its first edition in 2009. In that edition the assessment only made a reference to reading (Electronic Reading Assessment). In 2012, besides ERA, they have included computer based assessments of maths and problem solving. [2]

In the assessment on paper, in this edition 65 countries have participated. In Spain just over 25,000 students have been assessed, with expanded representative samples for fourteen Autonomous Regions. However, in the computer based assessment 19 countries have participated (16 from the OECD). In Spain, the database that allows the analysis of the data that comes from this type of assessment includes 10,175 individuals.

Also, among the collected variables, besides the scores, there are a series of indicators related to the supply and consumption of ICTs (Information and Communication Technologies) by the assessed students. This chapter considers the potential effects that the use and consumption of these new technologies can have as determinants on the competences assessed by computer in PISA 2012. This analysis is developed from a segmented view according to gender and partially extends to the type of the school.

This chapter is structured in different sections: Firstly, the Spanish situation is analyzed in relation to the knowledge of ICT's; secondly, the main empirical studies (international and national) on the impact of ICT on students' academic performance are examined; in the third, sample data and methodology used to develop the study is presented; the fourth shows the results; and finally the conclusions are presented.

SPANISH SITUATION WITH REGARD TO THE MASTERY OF INFORMATION AND COMMUNICATION TECHNOLOGIES (ICT)

From the information provided by EUROSTAT we can have a general idea of the actual use of ICT in Spain. Firstly, and for the latest available data (2013), it confirms that 70% of Spanish households had Internet access, 9 percentage points below the average for the EU-

[2] The section on problem solving was an area in PISA assessment in 2003 (on paper) but was reintroduced in subsequent editions of this publication to PISA 2012 (OECD, 2012).

28 and at a considerable distance from other countries such as France (82%), Germany (88%) and the UK (88%). However, this situation should not obscure the significant growth in internet coverage that has occurred in recent years. Secondly, 29% of the population accesses internet in order to search for information related to learning, while the European average stands at 32%. Finally, the better relative position of Spain in terms of the ability to use a computer and the internet must be considered. In this framework, the Spanish government approved in 2009 the 2011-2015 Strategy of the *Plan Avanza 2* (Advance Plan 2), which is divided into four lines: paperless administration, infrastructure, use and confidence in internet and the Spanish ICT industry drive.

With this framework, a closer approximation to the target population of our study is provided by the Survey on Equipment and Use of IT in households 2013 conducted by the INE (the Spanish Institute of Statistics). This survey shows that out of the total of the households, 73.4% had a computer and 69.8% had Internet access. However, there are broad differences depending on the socioeconomic background of the families. For example, with respect to having an Internet connection there is a wide variation from 40.5%, in the case of households with lower income to 97.1% in those households in the highest income level.

Regarding the use of the Internet, if we focus on the age range of 16 to 24 years old, a percentage higher than 90% of respondents indicated the use of email and participation in social networks. It is worth highlighting that 76.1% mentioned that they used Internet to search for information about education, training or other courses. Also, 80% said they access 'wikis' (like Wikipedia, for example) or on-line encyclopedias to gain knowledge on whatever subject. Finally, it should be noted that, in these four aspects, individuals in this age group show the highest rates of use in comparison to the other cohorts.

Also, the INE survey allows us to discover some of the habits related to ICT for the age group 10 to 15 years old, the closest group to the PISA assessed group. Here, a large majority should be considered as "frequent" computer and Internet users (95.2% and 91.8%, respectively), with marked differences by socioeconomic background but without notable differences by gender. Finally, as to where these children access the Internet, the survey shows that households are, with 88.2%, the most repeated option followed by schools with 70%.

With the aim of finding out the use of ICT in schools, we must turn to statistics provided by the Ministry of Education, Culture and Sports (MECD) via the General Department of Statistics and Research, whose latest version corresponds to the year 2011/2012. There we can see that, for all schools, 50.4% of computers were located in the classrooms, followed by 25.7% which were in the IT rooms. Specifically, in the case of public secondary schools, the proportion was 46.5% and 25.7% respectively. As for the use of ICT, these sources indicate that 77.7% of computers were intended for teaching activities, followed by 14% for teachers' own tasks.

Finally, an element that could directly affect the secondary students' knowledge of new technologies is the real possibility for them to have a computer for individual use at school. The sources of MECD indicate an average of 3.2 pupils per computer for teaching and learning tasks, 2.8 being the value in the case of public secondary schools.

THE IMPACT OF ICT ON EDUCATIONAL PERFORMANCE

The main studies that analyze the effect of the presence and use of information and communication technologies (ICT) in schools on students' performance (understood as passing tests, school course years or acquiring skills, for example) are not conclusive. So while one group of authors points out that ICTs have a positive impact on this performance, another group shows that they have no effect (and to a lesser extent that they have a negative impact).

Most studies analyze the implementation of specific policies that promote the presence and use of ICT. In them, the authors usually point out that the design and implementation (application) of the developed policies is often crucial. Other types of research link the presence of ICT resources in schools and their use on the results of the international assessments of competence acquisition (mainly PISA).

Among the group of analyses that show positive effects of the investment in ICT on school performance we can highlight the following: Machin et al. (2007) in England, Barrow et al. (2009) in the United States and Banerjee et al. (2007) in India. The study of Machin et al. (2007) shows that, among primary school students, an increase of the spending on ICT improves results in certain subjects (English and science) but not in maths. Also, Barrow et al. (2009) show the success of a computer program for maths learning in several districts of Los Angeles, so that in their study, students (randomly selected) supported by a computer performed better than those who followed traditional teaching methods. Finally, Banerjee et al. (2007) emphasize the positive effect of the use of a program for computer-assisted maths learning on the students' academic performance in urban schools in India. In this research the results of this program were compared against the results of hiring young teachers. Although both methods improved the students' performance, they were higher in the first case (with the use of computers). Moreover, the greatest benefits of the program were in those students who previously had low academic performance.

Among the studies of the second group, which mostly show the lack of impact of ICT on academic performance (and to a lesser extent a negative impact), we can highlight the following. Leuven et al. (2007) evaluate an assistance program to purchase computers and software in disadvantaged primary schools in the Netherlands. In their analysis there is no proof of any effect of the use of ICT on the students' scores on language and maths. Goolsbee and Guryan (2006) also analyze the grants given to school districts for schools to invest in internet and communications in USA. They observe that this investment has no significant impact in any of the analysed measurements related to academic performance. Rouse and Krueger (2004), in their evaluation of the effects of an educational computer program on knowledge in language and reading, do not find satisfactory results for the whole group of students, nor for those with learning difficulties. In Colombia, Barrera-Osorio and Linden (2009) reveal that the program to introduce computers to schools has not improved students' performance in maths and language, although the lack of success is associated with the use of the computers itself, as they dedicated more time to teaching the students how to use the computers than to teaching the subjects. Finally, in an evaluation of a program of resource provision for acquiring computers in schools providing primary and secondary education in Israel, Angrist and Lavy (2002) conclude that these resources do not affect the performance of students in maths in secondary education, and it even has a negative effect on the scores of students in primary education.

As we previously indicated, beyond the analysis of programs of incorporation and use of ICT for schools, there are studies that relate the provision and use of ICTs in schools with the acquisition of competences (assessed in international tests such as the PISA analysis of the OECD). The results are very different. Thus, for the 2000 edition, Woessman Fuchs (2004) show that computer use has no impact on students' performance. However, in the 2006 assessment, Spiezia (2010) points out that there is positive and significant effect of the use of computers in school on students' performance in science, and also this effect is greater in students with a higher socioeconomic level. On the other hand, the use of computers at home has more influence than the use of the computers at school.

In the Spanish context, there are no studies on specific policies, but the effect of ICT on students' performance is analyzed in international assessments. Thus, studies of Calero and Escardíbul (2007) and Cordero et al. (2012), using data from different PISA assessments, show no influence of the ratio of computers per student in academic performance. Choi and Calero (2013) suggest the existence of a negative effect. However, a recent analysis of Cabras and Tena (2013), using data from the PISA 2012 assessment, show some evidence of the positive effect of the use of computers on the students' acquisition of maths competences. In their study, this influence is significantly higher in students from more disadvantaged socio-economic groups. Finally, in their analysis of the assessment of different competences conducted by the Autonomic Region of Madrid between 2006 and 2009, Anghel and Cabrales (2010) conclude that the ICT variable is not significant for explaining the students' results. In summary, the research presents different results, although in the Spanish case the empirical evidence of the positive effect of ICT in school on the students' performance is barely proved.

Another element to bear in mind is the possible variation in the results of the assessments because of the tool used for its resolution. In this sense, Butters and Walstad (2011) use two groups of eighth and ninth grade students from schools in Florida and Delaware (13-15 years old) and find that students perform better on the computer based assessment than in an assessment on paper of equal difficulty. Also, while the computer based assessment reduces random answers, the assessment on paper reduces the bias caused by the order of the items assessed. In another study, Bennett et al. (2008) consider eighth graders (13-14 years) from public and private schools in the U.S., to determine the impact of the type of test on the results. They randomly form two groups of students to assess the same maths contents with computed-based and on-paper assessments, respectively. The results indicate a slight difference, but significant, in favor of assessments on paper. These same results are observed in the study of Jeong (2012) for South Korea. Meanwhile, Kingston (2009) in his extensive review of the literature for 1997-2007 and from a meta-analysis, concludes that the modes of assessments (paper or computer) do not appreciably affect the students' performance for the different analyzed grades for the U.S. One measurement to corroborate the comparability of the tests is introduced by Bennett (2003) who suggests that in order to confirm that the results obtained by both types of tests are comparable, the distribution of the obtained scores and the order among individuals between the scores from the computer-based and on-paper assessments should be maintained.

In fact, in the PISA-2012 assessment, in Spain the results are higher for students in the on-paper tests than for the computer-based tests. Thus, the maths result obtained in the computer-based assessment (475) is 9 points below the results obtained in the test on paper. For reading literacy, the difference in favor of the assessment on paper is 22 points (the result by computer is 466).

As for the possible differential effects by population subgroups, Kingston (2009) focuses attention on variables like the experience in the use of the tool, socioeconomic status, ethnicity and gender, without finding any significant differences. The study of Sandene et al. (2008) is along the same lines. Specifically, in terms of gender, Jeong (2012) finds that while women show a drop in performance when changing from paper to computer-based assessments in all the assessed competences, in the case of men this drop occurs only in language. In this regard, the PISA 2012 assessment shows that both boys and girls perform worse in computer-based assessments than on paper. The differences for boys are 11 points in maths and 21 in reading comprehension; in the case of girls these differences are, respectively, 7 and 23.

Regarding the effect of the hardware used (size of monitor, resolution), Květon et al. (2007) shows how a change in the colors of the test presentation affects the final scores. Bennett (2003) also focuses on the relevance of monitor size and its relation to the size of the text that can change the information shown on the screen to the student and therefore the subsequent performance. Other elements mentioned by Kingston (2009) are the quality of the Internet connection, the time given for the completion of the test and the students' preferences. In line with this last point, Richardson et al. (2002) indicates the importance of student's motivation as an element relevant for their test performance.

An additional element of analysis that the literature has studied is the influence of the potential limitations of the computer-based tests (being unable to review all the test questions, skipping questions or changing answers, etc.) in the obtained scores. In this line, a study by Mason et al. (2001) concludes that if the test conditions are equivalent (between computer-based and on-paper assessments) there are no differences between the scores obtained by the examinees. Therefore, both would be equivalent and would demonstrate the actual competence level of a student. Finally, Bennet (2003) adds that the computer-based assessments not only evaluate the abilities in a certain area (maths, science or language, for example) but also the skill to manage a different environment that requires additional competences. This last point is relevant in the study of Sandene et al. (2008) after conducting a comparative analysis of the computer-based and on-paper maths tests. The PISA 2012 assessments attempt to take this possible effect into account, explaining clearly that only one basic skills level for working with the computer is incorporated (Ministry of Education, Culture and Sport, 2013).

From the above, it is of particular interest to develop two types of studies. On the one hand, the effect of ICTs on the students' acquisition of competences in different areas. On the other, to see if these variables influence differently depending on whether the tests are completed on paper or using the computer. In this study we try to give an answer for the first proposed analysis. For the second, the study of Marcenaro (2014) in this volume can be reviewed.

DATA AND METHODOLOGY

From the database in PISA 2012 for computer-based assessments, with 10,175 observations, in this section, given its specific interest for this chapter, we explain in depth those variables that were considered in order to reflect the possession or use and consumption of information and communication technologies.

ICT variables

Most ICT-related variables used in this analysis are a scaled index using the "item response theory", in a way that the variables are transformed so that, in all OECD countries they have an average value of zero and a standard deviation equal to one (see OECD, 2013). A higher value of the index shows a greater use of the resource being analyzed or greater possession of elements associated with them.

The characteristics of the ICT variables used in the empirical analysis are described in Table 5.1. These variables can be considered in three groups: personal (which define the relationship of the students with ICTs); variables of possession of elements related to ICTs; and finally those concerning the use of ICTs. In the first case we consider both the attitude towards computers (whether for the students they are a tool to learn at school) as well as the age when they started using ICTs. In the second, we take into account the existence and availability of ICT in the household, the availability of ICT in schools and the ratio of computers to the number of students in the course year with most students aged 15 years old. It should be pointed out that in this group the variable indicating the availability of ICT at home and the variable that defines the existence of ICT resources in it are separately included in the same regression, since some repeated factors are incorporated in both. Finally, variables related to the use of ICT are added, whether they are used for entertainment, to do homework, to use at school (and the particular case of maths classes), as well as the time spent using computers. The main descriptive factors for the ICT variables of the database are shown in the Appendix, without detecting any significant differences according to gender.

Other variables

The control variables used refer to the characteristics of students and their families as well as schools. The first type of information is provided by the students, while the data referring to school variables come from the information given by the heads of the schools. The individual variables are: age of the student, whether the student is native or immigrant (first or second generation), whether the student has attended preschool education for over a year (compared to less than a year or not having attended), whether the student has repeated a course year (distinguishing between primary, secondary, or both) and the degree of absenteeism. Family variables include the following factors: whether the mother and father are in active employment, existence of more than 100 books at home, index of occupational status of the father and mother, whether it is a one-parent family and the years of schooling of each parent.

Table 5.1. Definition of the ICT variables used in the empirical analysis

Variable	Definition
Personal Variables related to ICT	
Attitude towards computers: Learning tool	Degree of agreement (strongly agree, agree, disagree, strongly disagree) with the following statements: the computer is a very useful tool for homework, homework with a computer is more fun, the Internet is a great resource for information that I can use for doing my homework; it is a problem using the computer for learning; as anyone can upload information to the Internet, in general it is not appropriate for school work, the information obtained via the Internet is generally too unreliable to use in homework
Age of starting to use ICT	Age in which the student used a computer for the first time
Existence of ICT	
Availability of ICT in the home	Availability at home for student's use of the following items: computer, laptop, internet, game console, mobile, MP3, MP4 player or similar; printer, USB memory
ICT resources in the home	There is educational software at home: there is internet connection; computers (the quantity should be indicated)
Availability of ICT at school	Availability in the school for the student to use the following elements: computer, laptop, internet connection, printer, USB memory
Computer/number of students Ratio	Number of computers available at school for education compared to the number of students (in the national modal course of 15 years old)
ICT use	
ICT use for entertainment	Union of various responses (never or almost never, once or twice a month, once or twice a week, almost every day, every day) to the following questions related to the use of computers for entertainment: games (single or group), use of email, chat, participation in social networks, watching videos, read news online, get practical information online, upload their own content to share.
ICT at home to do homework	Frequency (never or almost never, once or twice a month, once or twice a week, almost every day, every day) in which the student performs the following school work: browse the Internet for school work, use e-mail to communicate with classmates and teachers; download or upload material to the school website, check the school website for information, doing homework on the computer, to share learning materials with other students
ICT use at school	Frequency of computer use (never or almost never, once or twice a month, once or twice a week, almost every day, every day) for the following school activities: chat, use email; surf the internet for homework, download or upload school material from the Internet; upload school work on the school webpage, run simulations, and do excercises and practice (in foreign language or maths); doing homework on a school computer, use the school computers for group work and to communicate with other students
ICT use in maths	ICT use in maths class
Computer usage time	Computer usage time (in minutes on a typical day)

The school factors included in the analysis are: school type (public, semi-private, private-independent), location (municipality size in terms of the number of inhabitants), students groups by academic level (in some or all classes), autonomy in the administration of the school (considering teacher recruitment, wage setting and salary increases, establishment and allocation of the budget, school management and curriculum management), percentage of immigrant students, average educational level of the parents of students at the school, number of students per classroom, number of students per teacher, percentage of girls in the school, school size (number of students) and an index of quality of the teaching staff. Dummy variables indicating the Autonomous Region where the school is located are also included (although the database does not allow identification of what region is being referred to).

Multilevel hierarchical models and treatment of missing values

In the PISA assessments the data related to students is "nested" at a school level, so that simple linear regression techniques cannot be applied, since the selection of the students is random, but it is carried out once the schools have been selected. Therefore, the overall distribution of students is not random (a higher level of similarity between students from the same school can be expected). In this context, the appropriate econometric strategy is the use of hierarchical models (multilevel) which distinguish two categories: students (first level) and schools (second level).

Additionally, the assessment results are obtained using the methodology called "Item response theory" (IRT), which allows us to compare the results obtained by each student even though not all students answer the same questions. Specifically, in PISA five plausible values are provided as results for each student in each area of analysis. This entails that in the calculation of the estimators the five values should be used in a specific way: the statistical calculations for each plausible value are carried out and subsequently the average of the resulting values is considered. In reality, five plausible values and 80 replicas provided by PISA, which allow us to obtain efficient estimators, are used for the calculation of the statistics. The use of replicas is necessary due to the selection mode in two stages or levels of the sample described above.

The explanatory variables refer to students and schools. With respect to the former, the assessed students provide information related to family and personal issues. Regarding the schools, the heads of the schools report on the characteristics of the schools (material and staff) as well as on methods of management of schools and the processes of teaching-learning. There are three types of explanatory variables: quantitative, which come from questionnaire answers; indexes resulting from the transformation of one or more variables; and scaled indexes, using IRT methodology.

After completing the above, the multilevel analysis model estimated in this study is displayed in the equations (1) to (3):

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^{n} \beta_{1j} X_{kij} + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0,\sigma^2) \qquad (1)$$

$$\beta_{0j} = \gamma_{00} + \sum_{1} \gamma_{0l} Z_{lj} + \mu_{0j} \qquad \mu_{0j} \sim N(0,\tau_0) \qquad (2)$$

$$\beta_{1j} = \gamma_{10} \qquad \mu_{1j} \sim N(0,\tau_1) \qquad (3)$$

Where Yij refers to the student's expected maths score "i" in school "j"; Xkij is a vector of "k" characteristics of student "i" in school "j" (explanatory variables of level 1); Zlj is a vector of "l" characteristics of school "j" (level 2 variables). The random effects are μ_j (at school level) and ε_{ij}, (at student level). The estimated parameters are reported as β. Equation (4) allows us to present the entire model in a single equation; it is obtained by introducing equations (2) and (3) in (1). Thus, a series of fixed effects is distinguished from the random or stochastic effects .

$$Y_{ij}= \gamma_{00} + \gamma_{10} X_{kij}+\gamma_{0l} Z_{lj} +\mu_{0j} +\varepsilon_{ij} \qquad (4)$$

Also, in the empirical analysis the problem of the lack of responses of individuals in some variables (*missings*) is addressed. In this regard, the missing values were estimated using the imputation by regression method recommended by the OECD (2008). However, the percentage of *missing* values is small (no more than 5% in most variables used, reaching a maximum of 9.8%). Only in the case of variables relating to the immigrant status of students, school type and population where the school is located, have missing values not been assigned.

Finally, all estimations were carried out with the HLM 6 software, which allows estimations that take into account all the elements mentioned above and provides standard errors of the robust coefficients of the regression (Willms and Smith, 2005). In this respect, no multicollinearity was observed.

RESULTS

The results in the computer-based assessments for the three assessed competences show better performance in problem solving (476.8 in average), followed by the results in maths (474.9) and reading literacy (466.2). As for differences according to gender, in problem solving and maths a higher scores for boys (1.5 and 12.5 points respectively) is observed. In the case of reading literacy, girls are the ones who have better performance (26.9 more on average). As an additional note, as has been pointed out in the previous section, both in maths as well as reading literacy assessments, the students (as a whole) do worse on computer-based assessments (9 and 22 points, respectively). According to gender, the boys get 11 points lower in maths and 21 in reading literacy, whereas these differences in girls are 7 and 23 points, respectively.

Results for the three tests analyzed according to gender

Next, the results referring to the impact of the variables related to ICT on the acquisition of the three competences assessed by computer in PISA 2012 are presented: maths (Table 5.2), problem solving (Table 5.3) and reading literacy (Table 5.4). The sample has been divided according to gender in the three cases.

With respect to the maths competence, the personal variables related to ICT show that both the attitude toward computers as a learning tool as well as the age at which they have started using ICT has an influence in the case of the girls but not on their male classmates. For the girls, the results in the maths assessment are increased if they have a positive attitude towards computers as a work tool and the earlier they have started using them.

Table 5.2. Effect of ICT on the acquisition of competences in mathematics

Variables	Boys1	Girls1	Boys2	Girls2
Attitude towards computers	1.51 (1.40)	2.80** (1.31)	1.74 (1.42)	2.97** (1.30)
Age of starting to use ICT	-0.26 (0.55)	-1.41*** (0.52)	-0.28 (0.55)	-1.49*** (0.51)
ICT availability at home	2.71* (1.47)	1.89 (1.87)		
ICT resources at home			5.93*** (1.61)	4.38*** (1.54)
Availability of ICT at school	3.55** (1.44)	4.33*** (1.57)	3.93** (1.44)	4.92*** (1.57)
Computer/number of students	7.64 (6.34)	8.56* (4.96)	7.64 (6.34)	8.88** (4.98)
ICT use for entertainment	3.21* (1.71)	3.44* (2.07)	3.18* (1.74)	3.52* (2.02)
Use of ICT for school work at home	-3.30** (1.64)	-3.15 (2.03)	-2.84* (1.62)	-2.72 (2.06)
Use of ICT at school	-4.33*** (1.56)	-5.41*** (2.02)	-4.39*** (1.56)	-5.52*** (2.01)
Use of ICT for maths	-0.97 (1.48)	0.31 (1.23)	-0.75 (1.48)	0.42 (1.25)
Time spent using computers	-0.04 (0.03)	-0.12*** (0.03)	-0.04 (0.03)	-0.13*** (0.03)
Included constant	Yes	Yes	Yes	Yes
Personal/family/school variables	Yes	Yes	Yes	Yes
Regions included	Yes	Yes	Yes	Yes
Nº observations	4,983	4,821	4,983	4,821

*** indicates significant variable at 1%; ** at 5%, * at 10%.
Source: Microdata from PISA-2012.

As for the factors related to the existence of ICT in the home, from the two variables considered the one that shows higher significance for boys and girls is that which refers to the existence of ICT resources that takes into account whether there is educational software, computer/s and internet connection in the household. The variable that includes more factors in its definition has less significance and only for boys. In the case of the school, the availability of resources positively affects the performance of all students (boys and girls). These results are similar to the case of the *resilient* students, as explained in the chapter of Cordero et al. (2014) included in this volume. Likewise, in the case of the female students, a higher ratio of computers per student also has a positive impact on their results.

Finally, the variables of the use of ICT in general as well as the variables of time spent using ICT show negative results on the acquisition of competences, except in the case of the use of computers as entertainment. To our understanding, the interpretation of these results is as follows: greater use of ICT for entertainment (proxy of their greater mastery and enjoyment) increases the results, in line with the partial positive effect of the attitude towards computers, so those more familiar with them obtain better results. However, the negative sign of the greater use of ICT in terms of school work, at school or at home (in this case only for boys) may reflect an inverse causality: those who get worse results use computers more for this type of tasks because they need more time for completing homework and study. In the particular case of maths we can also analyze whether a higher use of ICT in class for this subject affects the acquisition of competences. The results, however, do not show any statistical significance.

Finally the time spent using computers shows a negative effect (in line with the results of the frequency of use), although only in the case of girls. In the area of problem solving, as shown in Table 5.3, with respect to personal factors related to ICT the variable associated with the attitude toward computers as a learning tool is no longer significant in problem solving, while the age at which the student has started using ICT does negatively affect both girls as well as boys.

The variables related to the existence of ICT at home (in the two elements included in the study) and at school are not significant, nor is the ratio of computers per student in the school. Thus, unlike the case of maths, the existence of ICT does not seem relevant in the performance in the computer-based assessment of problem solving.

Finally, regarding the variables of ICT use, again the use of computers as entertainment is positive, but the use of ICT at home for school work, as well as the time spent using them (in the case of girls) is negative. As explained for mathematics, the negative sign of the use of ICT variable refers to the existence of inverse causality.

Table 5.3. Effect of ICT on the acquisition of competence in problem solving

Variables	Boys1	Girls1	Boys2	Girls2
Attitude towards computers	-0.76 (1.94)	-1.99 (1.91)	-0.82 (1.94)	-1.99 (1.89)
Age of starting to use ICT	-2.00** (0.87)	-1.99** (0.78)	-1.79** (0.83)	-1.97** (0.76)
ICT availability at the home	-4.09 (2.58)	-0.89 (2.32)		
ICT resources at the home			1.44 (2.98)	1.66 (2.31)
Availability of ICT at school	0.71 (2.28)	3.29 (2.23)	-0.04 (2.29)	3.36 (2.26)
Computer/number of students	0.98 (7.68)	9.06 (6.67)	1.12 (7.72)	9.13 (6.65)
ICT use for entertainment	14.32*** (2.68)	11.52*** (3.04)	13.50*** (2.56)	11.16*** (2.97)
Use of ICT for school work	-12.18*** (2.31)	-8.86*** (2.64)	-12.03*** (2.29)	-8.74 (2.75)
Use of ICT at school	-4.36 (2.65)	-4.52 (2.99)	-4.22 (2.63)	-4.58 (2.97)
Time spent using computers	0.004 (0.05)	-0.12*** (0.04)	-0.006 (0.05)	-0.13*** (0.04)
Included constant	Yes	Yes	Yes	Yes
Personal/family/school variables	Yes	Yes	Yes	Yes
Included regions	Yes	Yes	Yes	Yes
Nº observations	4,983	4,821	4,983	4,821

*** indicates significant variable at 1%; ** at 5%, * at 10%.
Source: Microdata from PISA-2012.

Finally we analyze the impact of ICT on the acquisition of competences in reading comprehension (Table 5.4). In this case, the consideration of computers as a work tool does not affect the acquisition of competences, while starting to use ICT at an early age favors this acquisition only in the girls' case.

The availability of ICT resources does not influence the result of the reading comprehension assessment (except with a negative sign in one of the two variables associated with the availability of ICT at home). The practically significant non-influence of such variables makes the results similar to those associated with problem solving. With regard to the variables of ICT use, as in the previous cases, their use for entertainment has a positive effect on the results. As in other analyzed competences, in general, variables regarding ICT use (and time spent using them) show a negative significance. In the case of reading comprehension the use of ICT at home for school work negatively affects boys and girls, the use of ICT in school only in the case of boys and the time spent using them for the girls.

Table 5.4. Effect of ICT on the acquisition of competences in reading comprehension

Variables	Boys1	Girls1	Boys2	Girls2
Attitude towards computers	2.77 (1.89)	-0.65 (1.70)	2.65 (1.88)	-0.80 (1.70)
Age of starting to use ICT	-0.83 (0.71)	-1.31** (0.66)	-0.43 (0.70)	-1.12* (0.65)
ICT availability at the home	-8.07*** (2.21)	-5.65** (2.29)		
ICT resources at the home			2.40 (2.28)	1.83 (1.78)
Availability of ICT at school	1.04 (1.69)	2.80 (2.09)	-0.42 (1.75)	2.53 (2.08)
Computer/number of students	0.09 (6.58)	2.39 (6.34)	0.39 (6.79)	2.38 (6.41)
ICT use for entertainment	8.23*** (2.26)	8.35*** (2.77)	6.65*** (2.17)	6.85** (2.76)
Use of ICT for school work	-7.56*** (2.02)	-6.42*** (2.19)	-7.24*** (2.01)	-6.22*** (2.25)
Use of ICT at school	-4.02* (2.13)	-2.51 (2.27)	-3.74* (2.15)	-2.73 (2.26)
Time spent using computers	0.06 (0.04)	-0.06* (0.03)	0.05 (0.04)	-0.07** (0.04)
Included constant	Yes	Yes	Yes	Yes
Personal/family/school variables	Yes	Yes	Yes	Yes
Included regions	Yes	Yes	Yes	Yes
N. observations	4,983	4,821	4,983	4,821

*** indicates significant variable at 1%; ** at 5%, * at 10%.
Source: Microdata from PISA-2012.

A summary of the results is shown in Table 5.5 (the sign of the effect of each factor on the acquisition of competence is shown only if it is statistically significant). As it can be seen in the table, among the personal variables, while the attitude towards computers as a work tool barely affects the acquisition of competences (only for girls in maths), starting to use ICT at an early age is positively significant in all competences for girls (and in problem solving for boys). The availability of ICT resources at home and at school only affects the case of the maths assessment positively. Finally, there appears to be an inverse causality in the frequency of use of ICT at home and at school in most competences and students (more in the case of the former than the latter). Also, the time spent using computers (in minutes) is negatively related to the acquisition of competences only for girls. However, the use of ICT as entertainment, which may reflect a greater predisposition for their use and a degree of mastery, has positive effect on the acquisition of maths competences, problem solving and reading comprehension in the computer-based assessments of PISA, 2012.

Table 5.5: Signs of the effects of ICT in the acquisition of assessed competences

Variables	Maths		Problem solving		Reading comprehension	
	Boys	Girls	Boys	Girls	Boys	Girls
Attitude towards computers		+				
Age of starting to use ICT		-	-	-		-
ICT availability at the home	+				-	-
ICT resources at the home	+	+				
Availability of ICT in the school	+	+				
Computer/number of students		+				
ICT use for entertainment	+	+	+	+	+	+
Use of ICT for school work at home	-		-	-	-	-
Use of ICT at school	-	-			-	
Time spent using computers		-		-		-

In conclusion, different results are worth highlighting. Firstly, ICT variables affect the maths assessment more than the rest of assessed competences. Secondly, the type of ICT element included in the analysis is not irrelevant, so that not all the variables have the same impact. In this regard, in general, the positive effect shown by the variables related to the existence of ICT (at home or at school) on any competence (like maths) does not correspond with the time spent using the ICT for school work (which has a negative sign). Thirdly, the enjoyment of using computers (as entertainment) is positively related to the results on the computer based assessment. Finally, there is a greater effect of this type of variables on the computer based assessments than in the traditional results referred to by the studies that analyze the effect of these types of variables on PISA tests carried out on paper (see studies using the PISA database in its on-paper assessment, reviewed in section on impact of ICT on academic performance).

With regard to the other variables, the results highlight the impact of personal and family variables on the acquisition of competences, as usual in this type of study, over the effects of school variables (results are available on request). To highlight some results, preschool education attendance for over a year, the family's occupational and cultural status or the existence of educational resources in the home have a positive impact in practically all of the regressions. Repeating a course, absenteeism or immigrant status have a negative impact.

Finally, the models shown explain about 40%-45% of the variance of the model as a whole, a slightly lower value than usual in this type of analysis that consider assessments on paper (closer to 50%).

Results by type of school

Below, the results of an extension of the analysis are shown: the differential effect of ICT according to the type of school students attend (public or private). Maths is considered since it is the area assessed in depth in PISA 2012 and because it is the area in which ICT variables have been shown to have more influence.

As confirmed in Table 5.6, the effect observed for the set of female students (results in maths assessment increase if they have a positive attitude towards computers as a work tool and the earlier they have started using computer) occurs only in those who attend public schools. As for the factors related to the existence of ICT, those associated with home and school in the boys' case are significant and with a positive sign. However, when separating students according to school, the results show that the former (existence of ICT in the home) is statistically significant in the case of students in public schools and the second (existence of ICT in the school) for those in private schools. In the girls' case, the availability of ICT in schools is significant for students in both types of schools, while the variable associated with availability of resources at home is only significant in the case of students in public schools in one of the two definitions of the variable (that which contains a smaller number of factors in its definition.) Also, the positive impact of a higher ratio of computers per student is again confirmed only in the case of girls but in private schools.

With regard to the variables of ICT use, their use for entertainment is only significant in a regression for boys in public schools (not anymore for girls). Also, the use of ICT at home for homework, that was determinant only in the case of boys, appears to be statistically significant for those in public schools (in one of the two analyses). As for the use of ICT in schools, a negative relationship is observed only for boys and girls in public schools. Furthermore, the use of ICT in maths class is not significant for any type of school (as in the general regression). Finally, the time spent using computers is still significant in the case of girls (of any school).

Thus, although the sample size (at a school level) demands that we use caution when considering the results, this first exploration of the results reveals significant differences when dividing the sample by type of school and points to a route for future analysis.

Table 5.6. Effect of ICT on the acquisition of maths competences by type of school

	Public schools				Private schools			
Variables	**Boys1**	**Girls1**	**Boys2**	**Girls2**	**Boys1**	**Girls1**	**Boys2**	**Girls2**
Attitude towards computers	1.45 (2.12)	5.20*** (1.65)	1.73 (2.14)	5.44*** (1.65)	1.54 (1.76)	-1.11 (2.16)	1.65 (1.75)	1.08 (2.15)
Age of starting to use ICT	-0.15 (0.67)	-1.23* (0.70)	-0.22 (0.67)	-1.38** (0.68)	-0.58 (1.03)	-1.32 (0.87)	-0.57 (1.01)	-1.32 (0.88)
Availability of ICT at home	3.44* (1.87)	2.30 (2.28)			1.76 (2.28)	1.10 (2.94)		
ICT resources at home			6.90*** (2.04)	4.43** (1.77)			4.05 (2.69)	3.95 (2.71)
Availability of ICT at school	2.01 (1.64)	4.13** (1.87)	2.47 (1.64)	4.80** (1.88)	5.78** (2.42)	4.68* (2.73)	6.11** (2.47)	5.19* (2.80)
Computers/number of students	3.39 (10.95)	0.69 (9.02)	2.98 (10.95)	1.03 (9.02)	4.50 (6.01)	10.25** (4.30)	4.67 (5.96)	10.46** (4.32)
ICT use for entertainment	3.39* (2.02)	2.58 (2.43)	3.33 (2.08)	2.65 (2.42)	3.02 (2.75)	6.20 (3.91)	3.07 (2.70)	6.28 (3.84)
ICT use for school work at home	-3.41* (2.04)	-1.88 (2.37)	-2.75 (2.08)	-1.56 (2.38)	-1.89 (2.68)	-6.47 (3.96)	-1.78 (2.67)	-5.96 (4.02)
ICT use at school	-4.86** (2.10)	-6.40*** (2.11)	-4.94** (2.10)	-6.32*** (2.13)	-3.36 (2.28)	-4.76 (3.89)	-3.40 (2.28)	-5.24 (3.83)
ICT use for maths	-0.63 (2.01)	0.68 (1.59)	-0.24 (2.01)	0.74 (1.62)	-1.48 (2.02)	0.13 (1.90)	-1.43 (2.05)	0.31 (1.92)
Time spent using computers	-0.05 (0.04)	-0.11*** (0.04)	-0.06 (0.04)	-0.12*** (0.04)	-0.04 (0.06)	-0.13*** (0.04)	-0.04 (0.06)	-0.14*** (0.05)
Included constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Personal/family/school variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nº observations	2,650	2,727	2,650	2,727	2,172	2,094	2,172	2,094

CONCLUSIONS

This study analyzes the influence of ICT components collected in PISA 2012 database on the performance achieved by students in the computer based competence tests. Furthermore, these results are analyzed taking into account gender and type of school. The analysis focuses on one type of ICT variables that we consider to be important given the design of the test. Also, their choice is based on the results of their effect shown by the analyzed literature.

The context of the analysis is that of a country where the number of homes with Internet access shows some shortcomings compared to other European countries. Although there has been an improvement in recent years, there is still a significant difference based on the socioeconomic level of households. Even with this comparative disadvantage, Spanish children show a "frequent" contact with ICT tools.

The results obtained using a multi-level hierarchical methodology for the three computer-based tests indicate that the provision and use of ICT heterogeneously affects the different groups into which we have divided the original sample. The different results between gender and type of school end up justifying the partial analysis of the tests.

Just to highlight some results, the ICT variables affect the maths assessment to a greater extent than the rest of the assessed competences. Secondly, the ICT element included in the analysis is not irrelevant, so that not all the variables have an affect in the same way. In this regard, in general, the positive effect shown by the variables associated with the existence of ICT (in the home or in school) in some competence (like maths), does not correspond with the time spent using them for school work (which has a negative sign). Thirdly, the enjoyment of using computers (as entertainment) is positively related to the results in the computer-based tests. Finally, a greater effect of this type of variables is observed in computer-based tests than the results traditionally referred to by the studies that analyze the effect of this type of variables in the on-paper PISA tests.

The results of ICT use (and the age of starting to use them) are consistent with other analyses for the Spanish case. However, the positive effect of the possession of ICT (at home or at school), confined to maths in our study, represents something new with respect to previous analyses with PISA data.

As for the results according to type of school (public or private), they reveal translation mechanisms of the use of the tools with the obtained results that could be divergent. The limited sample size does not allow us to draw relevant conclusions for decision making but certainly provides a clear line for work in the future given that the real debate, still unresolved, is on how the ICT should be implemented in schools so that not only a technological change occurs in the classroom, but also an improvement in the teaching-learning method and, consequently, in the acquisition of competences by the students. In this regard, our study confirms a greater number of significant ICT variables in public schools than in private schools. As for gender, the results show that a greater number of ICT-related factors affect the academic performance of girls than that of the boys. Therefore, the suggested segmentations, both by gender and by type of school, are relevant.

REFERENCES

ANGHEL, B. Y CABRALES, A. (2010). "The Determinants of Success in Primary Education in Spain". Documentos de trabajo (FEDEA), 20, 1-65.

ANGRIST, J. D., Y LAVY, V. (2002). "New evidence on classroom computers and pupil learning". Economic Journal, 112, 482, 735-765.

BANERJEE, A. Et. al. (2007). "Remedying education: Evidence from two randomized experiments in India". Quarterly Journal of Economics, 122, 3, 1235-1264.

BARRERA-OSORIO, F., Y LINDEN, L.L. (2009). "The use and misuse of computers in education: Evidence from a randomized experiment in Colombia". The World Bank Human Development Network Education Team, Policy Research Working Paper 4836.

BARROW, L. Et. al. (2009). "Technology's edge: The educational benefits of computer-aided instruction". American Economic Journal: Economic Policy, 1, 1, 52-74.

BENNETT, R. (2003), "Online Assessment and the Comparability of Score Meaning, Educational Testing Service, Princeton", Nueva Jersey, www.ets.org/Media/Research/pdf/RM-03-05-Bennett.pdf.

BENNETT, R., J. Et. al. (2008), "Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP". Journal of Technology, Learning, and Assessment, 6, 9.

BUTTERS, R., Y WALSTAD, W. (2011). "Computer Versus Paper Testing in Precollege Economics". The Journal of Economic Education, 42, 4, 366-374.

CABRAS, S. Y TENA, J. D. (2013). "Estimación del efecto causal del uso de ordenadores en los resultados de los estudiantes en el test PISA 2012". En: INEE (Ed.), PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Análisis secundario. Madrid. INEE.

CALERO, J., Y ESCARDÍBUL, J. O. (2007). "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA-2003". Hacienda Pública Española, 183, 4, 33-66.

CHOI, Á., Y CALERO, J. (2013). "Determinantes del riesgo de fracaso escolar en España en PISA-2009 y propuestas de reforma". Revista de Educación, 362, 562-593.

CORDERO, J. Et. al. (2012). "Análisis de los condicionantes del rendimiento educativo de los alumnos españoles en PISA 2009 mediante técnicas multinivel". Presupuesto y Gasto Público, 67, 2, 71-96.

CORDERO, J. Et. al. (2014). "Superando las barreras: factores determinantes del rendimiento en escuelas y estudiantes con un entorno desfavorable". En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

FUCHS, T., Y WOESSMANN, L. (2004). "Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school". Brussels Economic Review, 47, 3-4, 359-385.

GOOLSBEE, A., Y GURYAN, J. (2006). "The impact of internet subsidies in public schools". The Review of Economics and Statistics, 88, 2, 336-347.

JEONG, H. (2012). "A comparative study of scores on computer-based tests and paper-based tests". Behaviour & Information Technology, 1-13.

KINGSTON, N. M. (2009). "Comparability of computer- and paper-administered multiple choice test for K–12 populations: A synthesis". Applied Measurement in Education, 22, 22–27.

KVĚTON, P. Et. al. (2007). "Computer-based tests: the impact of test design and problem of equivalency". Computers in Human Behavior, 23, 1, 32-51.

LEUVEN, E. Et al. (2007). "The effect of extra funding for disadvantaged pupils on achievement". Review of Economics and Statistics, 89, 4, 721-736.

MACHIN, S. Et. al (2007). "New technology in schools: is there a payoff?" Economic Journal, 117, 522, 1145-1167.

MARCENARO, O. (2014). "Del lápiz al ordenador: ¿diferentes formas de evaluar las competencias del alumnado? En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MASON, B. Et. Al. (2001). "An examination of the equivalence between non-adaptive computer based and traditional testing", Journal of Education Computing Research, 24, 1, 29-39.

MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE, Instituto Nacional de Evaluación Educativa (2013). "Marcos y pruebas de evaluación de PISA 2012. Matemáticas, Lectura y Ciencias". Madrid: Subdirección General de Documentación y Publicaciones.

OECD (2008). "Handbook on constructing composite indicators. Methodology and user guide". Paris: Organization for Economic Co-operation and Development (OECD).

OECD (2012). "PISA 2009 Technical Report", PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264167872-en

OECD (2013). "PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy", OECD Publishing. http://dx.doi.org/10.1787/9789264190511-en

RICHARDSON M. Et. al. (2002). "Challenging Minds? Students' perceptions of computer-based World Class Tests of problem solving". Computers in Human Behavior, 18, 6, 633-649.

ROUSE, C. E., Y KRUEGER, A. B. (2004). "Putting computerized instruction to the test: a randomized evaluation of a 'scientifically-based' reading program". Economics of Education Review, 23, 4, 323-338.

SANDENE, B. Et. al. (2005). "Online Assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project", Research and Development Series (NCES 2005–457). www.nces.ed.gov/nationsreportcard/pdf/studies/2005457_1.pdf.

SPIEZIA, V. (2010). "Does Computer Use Increase Educational Achievements? Student-level Evidence from PISA". OECD Journal: Economic Studies, 1, 1-22.

WILLMS, J.D. Y SMITH, T. (2005). "A Manual for Conducting Analyses with Data from TIMMS and PISA", Report prepared for the UNESCO Institute for Statistics. http://www.unb.ca/crisp/pdf/Manual_TIMSS_PISA2005_0503.pdf.

Appendix: Descriptive factor of the ICT variables

Variable	Sample size	Minimum	Maximum	Average	Typical Deviation	Average boys	Average girls	Difference boys-girls
Attitude towards computers: learning tool	9,181	-2.899	1.305	0.085	0.934	0.109	0.062	0.047
Age of starting to use ICT	9,644	5.000	16.000	7.755	2.357	7.647	7.862	-0.215
Availability of ICT at home	9,741	-4.018	2.783	0.172	0.869	0.235	0.109	0.126
ICT resources at home	10,038	-3.160	1.150	-0.009	0.812	-0.038	0.021	-0.059
Availability of ICT at school	9,691	-2.804	2.826	-0.092	0.969	-0.069	-0.116	0.047
Computers/number of students	9,928	0.039	8.000	0.720	0.533	0.724	0.717	0.007
ICT use for entertainment	9,518	-3.975	4.432	-0.025	0.810	0.089	-0.139	0.228
ICT use for school work at home	9,388	-2.444	3.733	0.080	0.868	0.052	0.108	-0.056
ICT use at school	9,439	-1.610	4.109	0.278	0.868	0.307	0.251	0.056
ICT use for maths	9,211	-0.775	2.801	0.024	1.029	0.078	-0.030	0.108
Time spent using computers	9,632	0.000	206.000	55.627	40.805	53.278	57.972	-4.696

6. Determinant factors of performance in problem solving

Spain in international perspective

Ildefonso Méndez

Universidad de Murcia

Summary

The aim of this chapter is to analyze the relative position of the countries of the OECD and, in particular, of Spain, in problem solving in PISA 2012. Spain performs below the OECD average in student performance, the adverse differential in problem solving being greater than that registered in maths. This result is not explained by the provision of family and schooling determinants of Spanish students, since this is higher than the OECD average. This result is obtained not only by analyzing the average score but also by using all the information of the distribution of students by competence levels defined by the OECD. We also find that non-cognitive skills or personality traits play a fundamental role in explaining differences between students and countries both in problem solving and in maths. In particular, both the non-cognitive skills reported by the student as well as the preferred non-cognitive skills to be encouraged in a child, reported by those living in the country of residence of the student several decades before the PISA 2012 exam, have an relevant effect on their score, even after controlling for a broad set of student, family and school characteristics. This result opens up a new way of explaining the adverse relative position of Spain in international student performance, since the personality traits most valued by the Spanish society and education system are not those that increase the probability of academic success.

Keywords

PISA, problem solving, Spain, OECD, non-cognitive skills.

INTRODUCTION

In recent years there has been a growing interest in the academic community to discover the origin of international differences in student performance. This interest is based on studies showing that student performance in early and middle stages of the education system is a predictor of both the access to higher education and wages in adulthood at the individual level, as well as international differences in economic growth (Hanushek and Woessman, 2011).

The Programme for International Student Assessment (PISA), developed by the Organization for Economic Cooperation and Development (OECD), is one of the reference databases for this type of analysis. The 2012 edition of PISA focuses on the assessment of knowledge acquired by students aged 15 in maths and includes, unlike the previous four editions, a module in which students are asked, in computer-assisted questions [1] about strategies of problem solving that arise frequently in everyday life.

The educational system has to lay the foundations of logical and rational thinking that helps to solve everyday problems, but does not provide regulated knowledge linked to specific matters that may be of direct use to this task. In this regard, the evaluation of international differences in problem solving has a component of transverse or global evaluation of international higher education systems, which the analysis of international differences in performance in specific subjects such as language, science or math underlies.

In this chapter we contribute to the literature on this subject by analyzing the relative position of the OECD countries in problem solving in PISA 2012. We also study how much of the observed differences between countries are due to differences in the provision of determinant factors of student achievement. In this sense, we conduct a distributional analysis which, unlike conventional analyses focused on the average score of each country, uses information from the entire distribution of students in competence levels. This analysis is more efficient because it considers the information of the distribution of students by competences levels and not just one moment of that like the average. The distributional analysis uses recent contributions contained in Herrero and Villar (2013) and Herrero, Méndez and Villar (2014).

Next we analyze the role that characteristics of students, their families and schools have in determining their performance in problem solving. In this sense, we pay special attention to the role that non-cognitive skills reported by students play in achieving good academic results. The 2012 edition of PISA includes, for the first time, two modules of questions in which students have to show their degree of agreement, on a scale of one to five, with a series of statements that reflect different levels of perseverance and preference for solving complex problems. The inclusion of these measurements as explanatory variables is very satisfactory from the point of view of goodness of fit, although the causal content of the estimated correlation is clearly conditioned by the possible endogeneity of non-cognitive skills reported by the student .

---

[1] Marcerano (2014), in this volume, analyzes the effect that the passage "of the pen to the computer" has on the performance of students in PISA. Their results suggest that this change in the process of assessment affects student performance and that changes in performance depend on the socioeconomic background of the students and their experience in the use of information and communication technologies (ICTs). In this sense, Mediavilla and Escardívul (2014), also in this volume, discuss in detail the effect that the possession and use of ICT has on student achievement when this is assessed by computer. These results also suggest that a direct comparison of the results obtained in problem solving in PISA 2012 and 2003, an earlier edition which assessed this competence in the same country, is not appropriate, since in 2003 the competence was assessed without the help of computers.

Because of that, and in order to go one step further in the analysis of the effect of non-cognitive skills in the academic performance of students aged 15 in PISA, we turned to the World Values Survey (WVS). This survey asks representative samples of a broad set of countries, which include nearly all OECD countries that we analyze, about non-cognitive skills or personality traits they consider a priority to encourage in children. Respondents have to choose up to five qualities of a list of eleven. We use the first two waves of this survey, conducted in the mid to late eighties of the last century, to obtain a measurement of the most valued non-cognitive skills in the countries of residence of the students surveyed in PISA, almost a decade before they were born. The new variable reports the result of social preferences, existing educational and employment institutions, etc., and prevailing economic conditions in each country. In any case, it reflects the non-cognitive skills most valued by each society as a result of the conditions mentioned.

In the econometric analysis we find that the result, in both problem solving and maths, of students in PISA 2012 survey depends, after controlling for numerous characteristics of the student, family and school, on the non-cognitive skills identified as most relevant in their home country several decades earlier. We found that there are certain non-cognitive skills that promote the achievement of good educational results, while others discourage it. These results let us establish a new theory to explain the unfavorable relative position of Spain in the international rankings of educational performance and achievement. This new theory points out that the non-cognitive skills most valued by the Spanish society, and its institutions, are not exactly those that lead to educational and career success. WVS data confirm this hypothesis.

The rest of the chapter is organized as follows. The second section presents a summary of the methodologies used in the empirical analysis. The third section presents the results of applying the methodologies to data from PISA 2012. Then, in the fourth section, we analyze the implications of the results obtained in the previous section, and finally the sixth section presents the conclusions reached.

METHODOLOGY

In this chapter we use two methodologies. First, we analyze the determinants of student performance in the subject assessed, problem solving, and also in maths, from a linear regression model estimated by ordinary least squares. Specifically, we propose the following regression model:

$$y_{isc} = \beta_0 + \beta_1 X_{is} + \delta Z_c + \varepsilon_{isc}$$

Where y_{isc} is the grade in a particular subject of the student *i*, who attends school *s* living in country *c*, X_{is} is a set of variables that measure certain characteristics of the student, their family and the school he/she attends which will potentially affect their performance, Z_c is a set of indicators of country of residence included to pick up systematic differences in performance between students from different countries not previously picked up and finally, ε_{isc} is an error term that is assumed to be normally distributed. Standard errors are corrected for clustering at the country of residence, as the dependent variable varies for students but the explanatory variables in Z_c only vary at a national level.

We also consider a variant of this specification in which the fixed effects of a country are replaced by a variable that reflects the unequal relative importance of a set of non-cognitive skills in the countries analyzed.

The other methodology used in this chapter is less conventional since it is the result of a recent contribution contained in Herrero and Villar (2013). These authors raise the question of how to compare the relative performance of various groups, from the distribution of the units that make them up, in the categories of an ordered categorical variable. The authors show that it is possible to derive an evaluation function to report on the relative position of each group using all of the distributional information and not just the average, as is usually the case.

Specifically, Herrero and Villar (2013) obtain a synthetic measurement of the relative position of countries which measures the probability that an individual randomly drawn from a country belongs to a higher competence category than that of another individual randomly drawn from another country. By extending the comparison to all countries or groups included in the analysis, Herrero and Villar (2013) find that it is the dominant eigenvector, from a matrix that summarizes all possible comparisons by pairs of countries, which summarizes the relative positions of countries.

The procedure described in Herrero and Villar (2013) can be explained, briefly, in the following terms. We have groups *g*, countries in our case, whose relative performance we want to evaluate and to do so we have the distribution of its members (students) in the values of a categorical variable ordered in categories *s* which, in our case, are the seven competence levels that the OECD defines for the subject of problem solving. We use a_{ir} to denote the proportion of group members *i* in the category *r* and we use the matrix A to sum up the corresponding values for all groups and categories.

We can say that the group *i* dominates the group *j* if it is more likely that a randomly chosen member of group *i* occupies a higher position than a randomly chosen member of group *j*. So, if *pij* represents the probability of that event, we can, in terms of the competence levels of PISA, define this variable as:

$$p_{ij} = a_{i7}\left(a_{j6} + a_{j5} + a_{j4} + a_{j3} + a_{j2} + a_{j1}\right) + a_{i6}\left(a_{j5} + a_{j4} + a_{j3} + a_{j2} + a_{j1}\right) + .. + a_{i2}a_{j1}$$

In comparisons by pairs, the *pij*/*pji* ratio reports the relative advantage of group *i* over group *j*, so that when this ratio is greater than one we can say that the group *i* has an advantage over the group *j*, and vice versa. However, when there are more than two groups, as in our case, comparisons by pairs only provide part of the necessary information useful for defining the relative position of each country. Herrero and Villar (2013) show that for $g > 2$ it is possible to find a vector that reports on the relative position of each group taking into account all possible comparisons with the other groups. The vector in question is the dominant eigenvector associated with a Perron matrix defined in the article by combining the different *pij*.

The method described in Herrero and Villar (2013) assumes that the groups being compared are homogeneous in determinants of the compared categorical variable, so that the comparison between groups reports differences not due to different distributions of those characteristics. Herrero, Mendez and Villar (2014) extend this result to the case in which countries or, more generally, the compared groups are heterogeneous in determinants of the categorical variable. The method presented in Herrero, Mendez and Villar (2014) allows us to equalize the distribution of a set of determinant factors of the latent variable underlying the categorical variable, in our case academic performance, in all categories and groups included in the analysis. To do this, Herrero, Mendez and Villar (2014) estimate discrete choice models which efficiently summarize the differences in the distribution of the determinant factors between each group and category, on the one hand, and a representative sample of the population whose distribution of determinants is taken as a reference, in our case the whole of the OECD, on the other hand.

Subsequently, each observation in each category and group is weighted by a factor whose function is to increase the relative importance of these observations with determinant factors that are over-represented in the reference sample, that is, the one whose distribution of determinants we want to impose in all categories and groups. Also, the weighting factor reduces the relative weight of the observations characterized by determinants under-represented in the reference sample. Thus, the distribution of determinants in all groups and compared categories is equalized. Then the procedure defined originally in Herrero and Villar (2013) is applied.

The focus proposed in Herrero and Villar (2013) is more relevant than the simple comparison of average grades and so it is more efficient since it uses all the information in the distribution of students by competence levels. Also, the contribution of Herrero, Méndez and Villar (2014), applied to our analysis, allows us to identify the relative position of Spain and other OECD countries once we control for a set of characteristics of the students, their family and their school which affect their educational performance.

The comparison between the eigenvectors obtained using the two approaches, ie: the comparison of the unconditioned and conditioned eigenvectors with determinants of student achievement lets us determine what part of the differences in educational performance between countries is due to differences in the provision of determinants between countries.

RESULTS

This section is divided into two sections. In the first, we analyze the relative position of Spain in problem solving in the set of OECD countries. Subsequently, we analyze the determinants of student performance in problem solving.

Problem solving: Spain in international perspective

Table 6.1 shows, for each of the 28 OECD countries included in the database, the average grade and the distribution of students by levels of competence in the subject of problem solving. [2] It is worth noting that, for consistency with the subsequent econometric analysis, the sample analyzed in Table 6.1 is not the total sample available for each country, but that for which the control variables or determinants of student performance considered in this chapter are defined. [3]

[2] Chapter 1 describes in detail the competence levels defined by OECD for problem solving.
[3] Chapter 2 describes the average scores for countries in the total sample available for each country in PISA 2012.

Table 6.1. Average score in problem solving and distribution of students by competence levels. PISA 2012. OCDE countries[1]

	Competence level							
Country	1	2	3	4	5	6	Average score	Sample size
Germany	0.033	0.089	0.182	0.259	0.269	0.168	531.98	2264
Australia	0.037	0.102	0.194	0.278	0.232	0.157	524.81	2637
Austria	0.042	0.100	0.223	0.290	0.241	0.104	514.73	2067
Belgium	0.045	0.086	0.174	0.270	0.257	0.168	528.56	2699
Canada	0.038	0.094	0.199	0.277	0.241	0.150	524.65	5060
Chile	0.093	0.188	0.274	0.272	0.133	0.040	473.05	2647
Rep. Korea	0.018	0.037	0.114	0.252	0.312	0.268	563.49	1965
Denmark	0.065	0.139	0.265	0.280	0.183	0.068	493.07	2493
Slovakia	0.079	0.129	0.246	0.280	0.190	0.077	493.97	2737
Slovenia	0.108	0.199	0.266	0.233	0.143	0.052	469.53	2370
Spain	0.086	0.132	0.243	0.278	0.184	0.077	491.52	3554
United States	0.026	0.108	0.229	0.282	0.232	0.122	520.31	2846
Estonia	0.027	0.096	0.208	0.317	0.234	0.117	521.21	2408
Finland	0.043	0.098	0.212	0.289	0.231	0.127	518.43	2765
France	0.033	0.068	0.183	0.316	0.270	0.130	527.62	7732
Netherlands	0.041	0.107	0.206	0.279	0.242	0.126	519.14	2638
Hungary	0.119	0.159	0.263	0.261	0.144	0.054	473.40	11969
Ireland	0.047	0.123	0.247	0.295	0.195	0.093	504.33	2875
Israel	0.187	0.163	0.212	0.210	0.151	0.078	462.72	3743
Italy	0.043	0.097	0.213	0.288	0.251	0.106	515.82	4264
Japan	0.009	0.037	0.131	0.275	0.334	0.216	557.32	2374
Norway	0.060	0.106	0.225	0.269	0.227	0.112	510.64	5793
Poland	0.083	0.146	0.264	0.272	0.162	0.073	486.56	3463
Portugal	0.058	0.142	0.260	0.294	0.185	0.060	493.62	3183
United Kingdom	0.028	0.062	0.200	0.287	0.274	0.150	534.23	3253
Czech Rep.	0.024	0.079	0.172	0.290	0.273	0.161	534.86	2489
Sweden	0.054	0.121	0.234	0.288	0.208	0.096	505.90	3106
Turkey	0.082	0.239	0.339	0.219	0.107	0.014	459.53	2741
OECD	0.056	0.114	0.218	0.275	0.219	0.118	510.73	98135

Note: 1 This sample is used in the estimation of the determinant factors.

This consideration is particularly relevant in the case of non-cognitive skills reported by the student. The 2012 PISA report includes, for the first time, two blocks of five questions each in which the student has to show his/her degree of identification on a scale of 1-5 with a series of statements that attempt to measure the degree of perseverance and preference for complex problem solving, respectively. These questions were included in two out of three questionnaires administered and these were then assigned randomly to students. Therefore, the inclusion of these variables in the analysis means a significant loss of sample size but not a selection bias, since the allocation of questionnaires to students was conducted randomly.

We chose to include these new variables in the analysis given the importance that the recent literature (Heckman, 2011) on determinants of educational performance assigns to non-cognitive factors or personality traits. Later in this section we review this literature and we analyze the content of the new questions included in PISA 2012.

As already mentioned in previous chapters, the PISA report classifies students into seven competence categories, increasing in student qualification in the problem solving competence. However, we group students into six competence categories, as reflected in Table 6.1. Specifically, we considered the two highest levels together, those indicative of the highest student qualification. This redefinition is essential to carry out a conditioned analysis in a high set of characteristics of the student, their family and their environment by the small proportion of students in the highest levels of competence in the less developed countries of the OECD.

According to Table 6.1, Spain gets an average score lower than that of the set of OECD countries in the competence assessed, that is, in problem solving. By analyzing the distribution of students by competence levels we observe that the lower average score of Spain is explained by the higher proportion of Spanish students in the two lower competence levels and the lower concentration of them in the higher levels. Specifically, the proportion of Spanish students in the two categories of lower qualification is about 5 percentage points higher than that of the set of OECD countries, while Spain builds up an adverse differential close to 8 percentage points in the two levels of highest competence.

On ordering countries by average rating, Spain is ranked No. 22 in a total of 28 countries. Evidently, the relative position in the average score is clearly correlated with the relative weight of the tails of the distribution. Thus, the country with the highest average score, the Republic of South Korea, has more than a quarter of its students concentrated in the highest level of competence that we distinguish in Table 6.1. The corresponding percentage for Japan, with the second-best average score, is 21.6%. These figures are a long way from the 7.7% of students in Spain in the highest level of competence, and even further from the meager 1.4% of Turkish students who reach this level, the lowest percentage of those listed in Table 6.1.

There is a clear correlation between the level of development of the OECD countries, measured by per capita income in purchasing power parity, and the average score in the problem solving competence. However, the relationship between the two variables is far from being deterministic, since the unconditioned correlation coefficient is 0.49. This result suggests that there are other determinants of student performance in problem solving that are independent of the level of development in their country of residence.

An interesting exercise is to check whether the relative position of Spain and the other OECD countries differ substantially when we analyze the competence of their students in maths. For this purpose, Table 6.2 provides the same information as in Table 6.1 but for the subject of maths. The comparison of descriptive statistics shows that Spain is below the OECD average in the two subjects analyzed, although the adverse differential for Spain is higher in problem solving than in maths. In particular, this adverse differential for Spain in the average score is about 19 points in problem solving and 7 in maths.

Table 6.2. Average score in maths and distribution of students by competence levels. PISA 2012. OCDE countries[1]

	Competence level						
Country	1	2	3	4	5	Average score	Sample size
Germany	0.077	0.172	0.283	0.295	0.173	540.67	2264
Australia	0.137	0.260	0.293	0.205	0.105	507.40	2637
Austria	0.097	0.241	0.317	0.241	0.104	518.16	2067
Belgium	0.087	0.183	0.276	0.271	0.183	537.77	2699
Canada	0.086	0.247	0.331	0.238	0.098	517.65	5060
Chile	0.334	0.301	0.235	0.109	0.021	451.95	2647
Rep. Korea	0.056	0.156	0.258	0.281	0.249	557.84	1965
Denmark	0.119	0.268	0.351	0.206	0.056	501.78	2493
Slovakia	0.188	0.258	0.272	0.195	0.087	493.98	2737
Slovenia	0.157	0.296	0.289	0.192	0.067	495.02	2370
Spain	0.125	0.253	0.340	0.230	0.053	502.01	3554
United States	0.177	0.294	0.291	0.175	0.063	490.82	2846
Estonia	0.062	0.220	0.359	0.256	0.103	526.49	2408
Finland	0.094	0.235	0.338	0.238	0.096	517.14	2765
France	0.111	0.228	0.310	0.243	0.108	516.07	7732
Netherlands	0.087	0.209	0.291	0.276	0.137	529.67	2638
Hungary	0.178	0.290	0.291	0.171	0.070	489.88	11969
Ireland	0.104	0.266	0.344	0.215	0.071	507.02	2875
Israel	0.230	0.268	0.268	0.171	0.062	479.96	3743
Italy	0.148	0.256	0.317	0.208	0.070	500.07	4264
Japan	0.068	0.168	0.293	0.279	0.192	544.53	2374
Norway	0.134	0.268	0.328	0.204	0.066	500.91	5793
Poland	0.088	0.240	0.309	0.234	0.128	523.31	3463
Portugal	0.180	0.263	0.286	0.201	0.069	492.79	3183
United Kingdom	0.104	0.251	0.308	0.231	0.106	514.20	3253
Czech Rep.	0.099	0.217	0.280	0.256	0.148	527.10	2489
Sweden	0.152	0.274	0.321	0.191	0.061	495.41	3106
Turkey	0.329	0.314	0.197	0.119	0.041	456.13	2741
OECD	0.132	0.247	0.303	0.220	0.099	509.02	98135

Note: 1 This sample is used in the estimation of the determinants.

The average scores in maths and problem solving of the countries included in the analysis are strongly correlated, the correlation coefficient being 0.73 between the two measurements. Later in this section we will analyze in detail the determinants of student performance in the two subjects.

The preceding analysis is inefficient in that it uses the average score used instead of the complete distribution of students by competence levels to reflect the relative position of each country. As highlighted in the previous section, Herrero and Villar (2013) analyze this problem and get a synthetic measurement of the relative position of countries which measures the probability that an individual randomly drawn from a country belongs to a competence category higher than another subject randomly drawn from any of the countries with which it is being compared. So Herrero, Méndez and Villar (2014) extend this result to the case where countries are heterogeneous in characteristics relevant for the compared variable. This contribution, applied to our analysis, allows us to identify the relative position of each country once we control for a set of characteristics of the student, their family and school that condition their educational achievement.

Comparing the eigenvectors obtained using the two approaches, ie: comparing the unconditioned and conditioned eigenvectors with determinants, we can determine which of the international differences in educational performance are explained by differences in the provision of determinant factors.

The characteristics of the student, their family and school included as determinants of their score in the conditioned eigenvector were selected from the economic literature on the subject (Hanushek and Woessman, 2011) and from an econometric analysis which verified its relevance, presented in detail later.

The school characteristics that we control in the analysis are: whether it is public or private, the size of the population center where it is located, whether it has other schools nearby, whether it has a shortage of teachers in subjects assessed in PISA, whether it has the ability to hire, fire or pay teachers, whether it groups students according to their performance in any subject, whether the scores of the students are used to analyze the effectiveness of teachers, the proportion of teachers with required qualifications and the average socioeconomic status of the students studying in the same school as the respondent.

The personal and family characteristics we consider are: age and sex of the student, whether they are first-generation immigrant, whether they are second-generation immigrant, educational level of parents and occupational status in the current, or in the previous job in case they are unemployed, a set of indicator variables on the number of books in the home, an indicator variable of whether the language usually spoken at home is the language of the country of residence or not and another variable which takes the value of one if the student attended classes in preschool school and zero otherwise.

We also control for the degree of perseverance and personal preference for problem solving reported by the students through their responses to ten questions included for the first time in the 2012 edition of PISA. In these questions, the student must indicate their degree of identification, on a scale of one to five, with five expressions that reflect a high or low level of perseverance and five other expressions that reflect high or low aptitude or preferences for complex problem solving. We redefine the student responses so that a higher value in the response indicates a higher stated level of perseverance and/or preference or aptitude for problem solving. Also, for each of the ten questions, we define an indicator variable that takes the value of one if the respondent reports feeling very identified or identified with the expression that denotes a high perseverance and willingness to solve problems.

Descriptive statistics of the characteristics of students and their families are presented in Table 6.3A. Table 6.3B describes the characteristics of the schools.

Table 6.3A. Descriptive statistics. Student and family characteristics. PISA 2012. OECD countries

	Student		Father's studies		Mother's studies		Father, occupational category							Mother, occupational category						
Country	Age	Woman	Higher	Int.	Higher	Int.	1	2	3	4	5	6 [a]	8	1	2	3	4	5	6 [a]	8
Germany	15.82	0.50	0.45	0.50	0.53	0.42	0.15	0.15	0.10	0.02	0.07	0.07	0.22	0.08	0.26	0.15	0.09	0.17	0.01	0.18
Australia	15.71	0.53	0.19	0.81	0.24	0.75	0.12	0.09	0.08	0.02	0.07	0.13	0.05	0.06	0.19	0.11	0.04	0.20	0.04	0.17
Austria	15.71	0.51	0.46	0.53	0.50	0.49	0.22	0.21	0.08	0.02	0.08	0.07	0.11	0.12	0.25	0.08	0.12	0.25	0.01	0.14
Belgium	15.82	0.50	0.43	0.57	0.51	0.49	0.14	0.09	0.13	0.02	0.04	0.15	0.07	0.05	0.25	0.13	0.09	0.22	0.07	0.10
Canada	15.71	0.51	0.63	0.33	0.71	0.26	0.16	0.17	0.11	0.02	0.08	0.13	0.05	0.07	0.27	0.17	0.10	0.22	0.02	0.10
Chile	15.82	0.49	0.22	0.78	0.24	0.76	0.05	0.08	0.08	0.02	0.14	0.14	0.13	0.02	0.11	0.14	0.11	0.25	0.04	0.21
Rep. Korea	15.82	0.51	0.46	0.52	0.30	0.69	0.07	0.21	0.13	0.06	0.08	0.11	0.05	0.01	0.15	0.24	0.17	0.25	0.02	0.12
Denmark	15.81	0.51	0.42	0.57	0.28	0.71	0.12	0.14	0.13	0.05	0.09	0.08	0.12	0.04	0.15	0.16	0.13	0.24	0.01	0.18
Slovakia	15.74	0.51	0.28	0.72	0.29	0.71	0.05	0.11	0.18	0.03	0.10	0.12	0.11	0.03	0.18	0.15	0.17	0.21	0.05	0.11
Slovenia	15.86	0.53	0.43	0.56	0.48	0.51	0.15	0.17	0.12	0.03	0.09	0.07	0.10	0.05	0.19	0.19	0.10	0.21	0.01	0.19
Spain	15.79	0.48	0.51	0.49	0.51	0.49	0.07	0.14	0.13	0.06	0.10	0.09	0.22	0.01	0.18	0.07	0.17	0.26	0.02	0.19
United States	15.82	0.49	0.23	0.38	0.11	0.27	0.05	0.07	0.04	0.06	0.26	0.09	0.16	0.01	0.04	0.01	0.03	0.05	0.00	0.83
Estonia	15.73	0.52	0.52	0.45	0.63	0.35	0.11	0.17	0.17	0.02	0.10	0.11	0.09	0.05	0.33	0.15	0.06	0.25	0.02	0.10
Finland	15.74	0.53	0.31	0.68	0.38	0.61	0.07	0.10	0.07	0.02	0.11	0.15	0.13	0.04	0.16	0.14	0.10	0.20	0.04	0.16
France	15.79	0.49	0.38	0.56	0.44	0.52	0.18	0.17	0.07	0.02	0.07	0.09	0.13	0.11	0.28	0.12	0.09	0.16	0.01	0.15
Netherlands	15.70	0.56	0.51	0.46	0.50	0.47	0.20	0.21	0.12	0.02	0.07	0.10	0.13	0.06	0.32	0.12	0.08	0.13	0.01	0.24
Hungary	15.84	0.52	0.54	0.44	0.63	0.36	0.12	0.17	0.10	0.02	0.06	0.09	0.17	0.09	0.27	0.13	0.11	0.18	0.02	0.16
Ireland	15.69	0.51	0.39	0.55	0.44	0.53	0.14	0.14	0.11	0.02	0.08	0.11	0.10	0.06	0.26	0.09	0.12	0.23	0.01	0.18
Israel	15.80	0.52	0.42	0.51	0.39	0.54	0.15	0.16	0.08	0.02	0.10	0.12	0.13	0.07	0.18	0.05	0.09	0.17	0.00	0.36
Italy	15.84	0.52	0.50	0.46	0.56	0.41	0.14	0.18	0.12	0.06	0.06	0.09	0.11	0.06	0.23	0.14	0.12	0.18	0.01	0.21
Japan	15.70	0.49	0.49	0.44	0.45	0.50	0.16	0.21	0.15	0.04	0.09	0.07	0.06	0.05	0.19	0.22	0.13	0.22	0.00	0.09
Norway	15.87	0.51	0.46	0.43	0.46	0.46	0.11	0.16	0.10	0.04	0.13	0.09	0.06	0.04	0.19	0.13	0.08	0.22	0.01	0.26
Poland	15.78	0.52	0.45	0.52	0.57	0.40	0.13	0.18	0.08	0.02	0.15	0.07	0.12	0.06	0.23	0.12	0.09	0.26	0.01	0.16
Portugal	15.73	0.48	0.26	0.73	0.32	0.68	0.13	0.10	0.14	0.03	0.09	0.10	0.09	0.06	0.14	0.13	0.13	0.21	0.01	0.19
United Kingdom	15.75	0.51	0.19	0.44	0.22	0.47	0.09	0.10	0.08	0.04	0.13	0.13	0.13	0.04	0.14	0.05	0.10	0.24	0.08	0.27
Czech Rep.	15.79	0.50	0.51	0.48	0.59	0.39	0.11	0.23	0.10	0.03	0.12	0.07	0.11	0.05	0.43	0.11	0.08	0.19	0.01	0.09
Sweden	15.71	0.47	0.52	0.46	0.41	0.58	0.11	0.15	0.19	0.15	0.12	0.06	0.06	0.02	0.16	0.10	0.12	0.22	0.01	0.29
Turkey	15.76	0.49	0.25	0.71	0.29	0.68	0.02	0.11	0.15	0.07	0.15	0.09	0.06	0.01	0.17	0.11	0.12	0.21	0.01	0.29

Table 3A. (cont) Descriptive statistics. Student and family characteristics. PISA 2012. OECD countries

	Number of books in the home					Immigrant of generation		
Country	11-25	26-100	101-200	201-500	> 500	1	2	Language at home [b]
Germany	0.17	0.31	0.16	0.11	0.05	0.20	0.08	0.87
Australia	0.19	0.34	0.17	0.12	0.08	0.01	0.00	0.99
Austria	0.14	0.30	0.19	0.16	0.09	0.14	0.08	0.94
Belgium	0.13	0.32	0.21	0.18	0.09	0.18	0.02	0.95
Canada	0.14	0.35	0.20	0.16	0.06	0.15	0.10	0.84
Chile	0.18	0.35	0.17	0.09	0.05	0.04	0.01	0.94
Rep. Korea	0.11	0.28	0.23	0.21	0.12	0.12	0.02	0.95
Denmark	0.15	0.31	0.18	0.15	0.10	0.15	0.06	0.91
Slovakia	0.12	0.34	0.20	0.17	0.09	0.07	0.03	0.97
Slovenia	0.17	0.28	0.18	0.15	0.07	0.18	0.05	0.93
Spain	0.13	0.35	0.20	0.16	0.08	0.01	0.01	1.00
United States	0.28	0.27	0.10	0.06	0.03	0.02	0.01	0.95
Estonia	0.12	0.30	0.21	0.18	0.10	0.17	0.07	0.90
Finland	0.12	0.27	0.18	0.17	0.17	0.05	0.01	0.99
France	0.13	0.30	0.20	0.18	0.10	0.24	0.10	0.91
Netherlands	0.17	0.31	0.18	0.14	0.11	0.23	0.07	0.89
Hungary	0.14	0.31	0.21	0.16	0.08	0.18	0.09	0.83
Ireland	0.15	0.29	0.20	0.16	0.07	0.15	0.14	0.96
Israel	0.26	0.31	0.12	0.08	0.04	0.02	0.02	0.99
Italy	0.15	0.30	0.18	0.15	0.08	0.17	0.08	0.78
Japan	0.18	0.29	0.15	0.14	0.07	0.14	0.04	0.94
Norway	0.12	0.29	0.23	0.19	0.11	0.06	0.10	0.83
Poland	0.17	0.32	0.17	0.13	0.06	0.21	0.07	0.89
Portugal	0.22	0.34	0.15	0.08	0.05	0.13	0.03	0.94
United Kingdom	0.21	0.29	0.15	0.11	0.05	0.15	0.07	0.97
Czech Rep.	0.11	0.30	0.21	0.20	0.10	0.13	0.06	0.93
Sweden	0.08	0.26	0.23	0.26	0.13	0.00	0.00	1.00
Turkey	0.19	0.29	0.17	0.13	0.09	0.08	0.06	0.83

Note: The reference category in the regression for each of the variables is omitted. The occupational categories are: business management and government (1); technical and scientific and intellectual professionals (2); technicians and support professionals (3); administrative employees (4); workers in catering, personal, security and sales services (5); skilled agricultural and fishery workers and craftsmen and skilled workers (6, 7); machinery operators and assemblers (8); unskilled workers (9). [a] Occupational categories 6 and 7 were grouped to achieve a sufficient number of observations in each category. [b] Indicates whether the language spoken most of the time at student's home is the language of the country of residence.

Table 3B. Descriptive statistics. Characteristics of schools. PISA 2012. OECD countries

		City [a]		Lack of [b]		Autonomy of school [c]				Teachers	Schools	Students	Assesses
Country	Private	average	large	teachers	resources	Hire	Salary	Budget	subjects	qualified[d]	near[e]	grade[f]	teachers[g]
Germany	0.11	0.37	0.35	0.14	0.14	0.97	0.58	0.82	0.77	0.98	0.78	0.97	0.60
Australia	0.07	0.26	0.28	0.21	0.21	0.88	0.26	0.66	0.81	0.86	0.67	0.67	0.61
Austria	0.48	0.35	0.34	0.23	0.12	0.99	0.86	0.97	0.93	0.97	0.94	0.99	0.89
Belgium	0.15	0.30	0.35	0.25	0.26	0.88	0.30	0.72	0.75	0.87	0.75	0.80	0.58
Canada	0.13	0.35	0.33	0.26	0.25	0.78	0.24	0.64	0.69	0.76	0.65	0.78	0.48
Chile	0.11	0.40	0.23	0.18	0.44	0.86	0.38	0.68	0.73	0.81	0.73	0.74	0.58
Rep. Korea	0.24	0.35	0.36	0.36	0.14	0.71	0.15	0.43	0.65	0.92	0.86	0.78	0.47
Denmark	0.15	0.29	0.33	0.22	0.25	0.82	0.28	0.53	0.68	0.67	0.64	0.73	0.46
Slovakia	0.18	0.43	0.28	0.12	0.24	0.91	0.76	0.84	0.88	0.92	0.90	0.59	0.63
Slovenia	0.11	0.32	0.32	0.30	0.21	0.86	0.26	0.70	0.80	0.82	0.69	0.73	0.55
Spain	0.32	0.19	0.74	0.15	0.08	0.45	0.27	0.57	0.84	0.99	0.93	0.78	0.76
United States	0.10	0.30	0.39	0.30	0.31	0.60	0.20	0.64	0.49	0.77	0.68	0.76	0.61
Estonia	0.19	0.36	0.38	0.26	0.22	0.83	0.33	0.69	0.69	0.73	0.71	0.83	0.53
Finland	0.25	0.38	0.39	0.23	0.19	0.87	0.35	0.77	0.75	0.82	0.75	0.83	0.59
France	0.41	0.22	0.59	0.37	0.11	0.81	0.23	0.84	0.73	0.96	0.94	0.96	0.50
Netherlands	0.13	0.32	0.41	0.32	0.23	0.83	0.30	0.69	0.75	0.78	0.74	0.83	0.55
Hungary	0.12	0.23	0.45	0.19	0.13	0.86	0.17	0.60	0.54	0.94	0.75	0.93	0.34
Ireland	0.28	0.38	0.30	0.25	0.20	0.80	0.22	0.74	0.63	0.79	0.74	0.88	0.49
Israel	0.70	0.25	0.60	0.42	0.15	0.85	0.65	0.80	0.73	0.94	0.91	0.64	0.56
Italy	0.67	0.46	0.26	0.32	0.11	0.87	0.10	0.72	0.67	0.32	0.93	0.79	0.39
Japan	0.27	0.35	0.37	0.25	0.21	0.87	0.40	0.81	0.86	0.71	0.77	0.78	0.59
Norway	0.45	0.33	0.43	0.04	0.09	0.48	0.10	0.82	0.72	0.96	0.91	0.83	0.53
Poland	0.31	0.38	0.30	0.17	0.24	0.91	0.28	0.83	0.76	0.86	0.79	0.84	0.46
Portugal	0.05	0.50	0.32	0.03	0.24	0.96	0.22	0.75	0.76	0.89	0.77	0.71	0.36
United Kingdom	0.14	0.34	0.29	0.20	0.23	0.74	0.20	0.75	0.45	0.79	0.73	0.71	0.52
Czech Rep.	0.06	0.21	0.25	0.28	0.21	0.90	0.21	0.72	0.72	0.91	0.54	0.65	0.49
Sweden	0.37	0.18	0.71	0.15	0.15	0.45	0.08	0.50	0.80	0.99	0.85	0.92	0.78
Turkey	0.11	0.44	0.35	0.25	0.09	0.20	0.06	0.27	0.86	0.92	0.61	0.76	0.28

Note: [a]Indicates that the school is in a population center of between 15,000 and 100,000 people (average) or more than 100,000 people (large). [b]Indicates whether there is a shortage of qualified maths, science or language teachers or whether they lack computers in the school. [c] Indicates whether the school has the ability to hire teachers, determine their starting salary, decide budget allocation within the school or determine the contents of the subjects. [d] Reports on the percentage of teachers in the school with the required qualifications. [e] Reports on whether there are other schools near the school where the respondent studies. [f] Indicates whether students of the school where the respondent studies are grouped according to their performance in at least one subject. [g] Indicates whether the grades of the students are used to assess the performance of teachers.

Table 6.4 presents the eigenvectors unconditioned and conditioned to characteristics, and the average score of each country. Both the components of the eigenvectors as well as the average scores are presented as standardized to facilitate their interpretation. Specifically, the average score has been standardized making the average score for the set of OECD countries equal to one unit. Meanwhile, the normalization in the eigenvectors happens because it determines that the sum of the components equals the number of components of the eigenvector. Thus, higher unit values in one component of the eigenvector indicate that the country has a higher than average relative position, while values lower than the unit are typical of countries whose relative position is lower than that of the average of the countries considered.

A first result that comes from Table 6.4 is that the consideration of only the average score attenuates the differences between countries with respect to the alternative of considering the differences in distribution of students by competence levels[4]. Specifically, the coefficient of variation among OECD countries in the eigenvector not adjusted by characteristics is more than eight times higher than that obtained using the average score per country in problem solving. This result occurs because the average score overestimates the relative position of countries with results below the set of the OECD such as, for example, Chile, Slovenia, Hungary, Israel or Turkey, whose components in the eigenvector are approximately 45% lower than the relative position measured with the average score. A similar result in qualitative terms is obtained for Spain, whose position relative to the OECD average is 27% lower when we use all the distributional information, that is, according to the eigenvector, than when we use only the average score.

[4] This same result was obtained in Herrero, Méndez y Villar (2014) after analyzing the differences between countries of the OECD in student performance in language and maths.

Table 6.4. Average score and unconditioned and conditioned eigenvectors. Problem solving. PISA 2012. OECD countries

Country	Average score	Eigenvector Unconditioned	Eigenvector Conditioned
Germany	1.042	1.302	1.235
Australia	1.028	1.147	0.953
Austria	1.008	0.991	1.046
Belgium	1.035	1.266	1.720
Canada	1.027	1.152	0.871
Chile	0.926	0.506	0.569
Rep. Korea	1.103	2.248	1.790
Denmark	0.965	0.699	0.658
Slovakia	0.967	0.725	0.858
Slovenia	0.919	0.503	0.599
Spain	0.962	0.706	0.626
United States	1.019	1.035	1.041
Estonia	1.021	1.076	0.957
Finland	1.015	1.044	0.894
France	1.033	1.258	1.519
Netherlands	1.016	1.050	1.112
Hungary	0.927	0.539	0.535
Ireland	0.987	0.830	0.948
Israel	0.906	0.517	0.507
Italy	1.010	1.021	1.130
Japan	1.091	2.057	1.941
Norway	1.000	0.932	0.991
Poland	0.953	0.647	0.624
Portugal	0.966	0.696	0.970
United Kingdom	1.046	1.325	1.176
Czech Rep.	1.047	1.365	1.298
Sweden	0.991	0.854	0.716
Turkey	0.900	0.508	0.718
Coeff. of variation	0.051	0.418	0.374

Also, the average score underestimates the relative position of countries that get results higher than the OECD average. This is the case, among others, of South Korea and Japan, whose relative superiority is, when measured with the eigenvector, 104% and 88% higher, respectively, than when measured with the average score.

These results confirm the need to use all the distributional information available to characterize international differences in educational results. We also get differences in the relative position of the countries considered, though smaller, when we control for international differences in the provision of personal, family, and school characteristics that determine educational performance. Specifically, the differences between countries are reduced by 10.5% when we control for differences in these characteristics.

This result is a consequence of the fact that usually the countries that get better results in problem solving have an provision of characteristics of students, their families and their schools which are more favorable to the achievement of these results. For example, the educational level of the parents is usually higher in these countries, their jobs tend to be better paid and correspond to better occupational categories, there are more books at home, more reading and study habits, schools better equipped with human resources and materials, etc. Thus, the relative advantage of the Republic of South Korea, Australia and Canada is reduced by at least 17% when controlling for differences in characteristics, only the first of these countries maintaining its relative position above the OECD average.

Other countries that get unconditioned scores below those of the set of OECD markedly improve their relative position when we discount international differences in provision of personal, family and school resources. This is the case of Portugal and especially Turkey, whose relative disadvantage is reduced by approximately 40% when taking into account that their provision of determinant factors is particularly unfavorable to achieving good educational results.

The performance of Spain is less favorable, since it is, together with Sweden, the only country whose relative disadvantage widens considerably, meaning at least 10%, when controlling for characteristics. This result is indicative that Spain should, by its provision of factors, get better results than it does. In other words, and using a terminology more typical of a Oaxaca decomposition, it is the performance of the determinants considered in the analysis, not their provision, which helps explain the unfavorable relative position of Spain in problem solving in the context of the OECD countries.

Table 6.5. Average score and unconditioned and conditioned eigenvectors. Maths. PISA 2012. OECD countries

	Average	Eigenvector	
Country	score	Unconditioned	Conditioned
Germany	1.062	1.572	1.301
Australia	0.997	0.908	0.917
Austria	1.018	1.078	0.900
Belgium	1.056	1.488	2.416
Canada	1.017	1.075	0.863
Chile	0.888	0.366	0.372
Rep. Korea	1.096	1.983	1.739
Denmark	0.986	0.837	0.764
Slovakia	0.970	0.763	1.002
Slovenia	0.973	0.738	0.868
Spain	0.986	0.865	0.832
United States	0.964	0.682	0.684
Estonia	1.034	1.232	1.111
Finland	1.016	1.067	0.967
France	1.014	1.072	1.270
Netherlands	1.041	1.306	1.252
Hungary	0.962	0.693	0.730
Ireland	0.996	0.912	0.998
Israel	0.943	0.619	0.719
Italy	0.982	0.822	0.786
Japan	1.070	1.650	1.687
Norway	0.984	0.824	0.746
Poland	1.028	1.151	1.167
Portugal	0.968	0.745	0.860
United Kingdom	1.010	1.036	0.766
Czech Rep.	1.036	1.249	1.073
Sweden	0.973	0.760	0.583
Turkey	0.896	0.506	0.626
Coeff. of variation	0.047	0.404	0.357

Table 6.5 presents the same information as Table 6.4 for the subject of maths. The objective of this analysis is to help to discern whether the results for Spain are specific for the problem solving competence or, on the contrary, are rather general, country-specific and observed in other subjects. The results in maths confirm that the relative position of Spain is reduced when we consider the entire distribution of students by competence levels and not just the average maths score and also when we control for characteristics. However, both the reduction in relative position due to controlling for determinant factors and the resulting relative

disadvantage are lower in maths than in problem solving. This result confirms that Spain should, given its provision of student, family and school characteristics, get better results than those obtained in the assessed subjects.

Applying a traditional procedure of decomposition of observed average differences (Oaxaca) confirms this result by noting that all of the observed difference in average score between Spain and the OECD has its origin in differences in performance and not in differences in provision of determinant factors.

Spain has, compared to the set of the OECD, a higher proportion of fathers with intermediate/average studies, parents currently employed or in a previous employment in the first two occupational categories, a high proportion of households with a large number of books in the home, a greater proportion of students in schools that separate students according to performance and a higher proportion of students who report high levels of perseverance and preference for problem solving. However, the effect of having a mother with studies above primary is lower in Spain, as is the effect that the mother or father has a high occupational status or the effect that the grades of the students are used to determine the performance of the teacher.

The determinant factors of performance in problem solving

Next, we present a detailed analysis of the determinants of student performance in problem solving. In this analysis we use the technique of ordinary least squares regression to determine the relative importance of each of the individual, family and school characteristics listed above in a multivariate analysis.

Before discussing the results it is worth noting that we found a logical high correlation between the responses that a student provides for the same five questions about their level of perseverance, but also between these responses and those for five questions on problem solving. So in order to efficiently summarize the information contained in these questions and avoid problems of multi-collinearity, we obtain the first principal component of the answers to the ten questions and we label it as our synthetic measurement of non-cognitive skills reported by the student. Table 6.6 presents the weightings of the ten variables in obtaining the first principal component.

Non-cognitive skills or personality traits play a prominent role in current microeconometric analysis and particularly in the study of the determinants of educational performance. The review of the literature on the psychology of personality and economics held in Heckman (2011) and Almlund and others (2011) allows us to conclude that certain personality traits such as, for example, perseverance are the most relevant in explaining the differences observed in the adult population in education, career and health results (Conti, Heckman and Urzua, 2011). These studies highlight that non-cognitive skills have an importance at least equivalent to cognitive skills in explaining education, health and career success in adulthood. Skills such as perseverance, self-sacrifice, and hard work are particularly relevant in achieving those results.

Table 6.6. Relative importance of statements about perseverance and preference for solving problems in obtaining the first principal component. PISA 2012. OECD countries.

Statement	Factor
When faced with a problem, I give up straight away	0.254
I put off difficult problems	0.227
I stay interested in tasks that I start	0.287
I continue working on a task until everything is perfect	0.316
When faced with a problem, I do more than is expected of me	0.315
I can handle a lot of information at the same time	0.344
I understand things quickly	0.355
I look for the explanation for things	0.326
I can connect different pieces of information quickly	0.366
I like solving complex problems	0.342

A recurring example in the literature on the subject is the Perry preschool intervention program, implemented in the U.S. and directed towards students aged between 3 and 4 years old from disadvantaged socio-economic backgrounds. The program selection was done randomly and the treatment consisted of school support classes and sessions that encouraged self-control and other positive aspects of personality.

Heckman et al (2010) concludes that the Perry program improved the results of participating students in terms of educational level reached, employment, wages, participation in healthy activities and criminal behavior over 30 years after being implemented. This result is not explained by the effect of the program on the accumulation of regulated knowledge or cognitive aspects, since the differences in IQ between participants and non- participants were not statistically significant shortly after the program was implemented.

Heckman, Pinto and Savelyev (2012) show that the key to the effectiveness of the Perry program is that the program significantly increased the provision of favorable non-cognitive skills in participating students. Thus, these students achieved levels of self-control, perseverance and motivation, among other non-cognitive characteristics, which were significantly higher than those that they would have had they not participated in the program. Despite the fact that their IQ did not improve compared to non-participants, they also achieved consistently better results in their academic performance. The strength of these results and their long term effect placed non-cognitive skills at the center of economic analysis in general, and particularly microeconometric analysis.

Finally, the Education Department of the United States underlined the need to promote tenacity and perseverance as critical factors for educational success in the twenty-first century, in a recent report of February 2013.

For all these reasons, we think it is crucial to include the self-classification of students in the skills of perseverance and preference for problem solving in the set of explanatory variables of their performance.

Table 6.7 summarizes the results of estimating, by ordinary least squares, the grade obtained in problem solving on the set of features described above. The estimation was performed using the corresponding weighting factors and in accordance with the "statistical shortcut" described in OECD (2009). We present two estimations, one for the set of OECD countries and another for Spain. The estimation for the set of OECD countries includes variables indicating country to control for fixed effects of country.

A first result to highlight is that the personal and family characteristics of students are essential for explaining the differences observed in educational performance. Thus, the sex and age of the student, educational level, and above all, the occupational category of their parents, the migration history of the family or the number of books at home, have a statistically significant and quantitatively relevant correlation with student performance in problem solving.

The school characteristics are also important, although notably less in terms of goodness of fit. Specifically, public or private ownership of the school determines, in statistically significant terms, the grade in problem solving, at least for the set of OECD countries analyzed. We also found that having studied preschool education significantly increases the performance of students in problem solving. This result, unlike what was previously mentioned, is obtained both for the OECD as a whole and for Spain. This result, already highlighted internationally in OECD (2011) suggests that students who attend preschool education for at least one school year get better results, even after conditioning socioeconomic characteristics of their environment, than those who don't even attend one year of preschool education. This result is best interpreted from the contributions contained in Heckman (2008). This study, highlighted among others with similar results to Felfe, Nollenberger and Rodríguez -Planas (2012), shows that the performance, both private and social, of interventions leading to the reduction of inequality and the encouragement of both appropriate cognitive and non-cognitive skills in students is greater when these interventions occur at early ages. Thus, preschool education would be an example of early intervention of high social performance.

Also, there are two variables that report school characteristics with statistically significant coefficients but only in the estimation for Spain. These are the indicators of the presence of other schools in the vicinity and the indicator that the school separates students according to their performance in at least one subject. The first variable, that is the presence of other schools near the school attended by the student is negatively correlated with student performance in problem solving. However, we did find a positive correlation between attending a school in which students are separated according to their performance in at least one subject and the average performance of students.

The synthetic indicator variable of non-cognitive skills reported by the student shows a positive and highly significant correlation with the grade obtained in problem solving. This variable provides additional information to that contained in the specification that doesn't include it, since the goodness of fit of the model is increased by more than 10% after its incorporation as an explanatory variable. This result is interesting in that it confirms, for the set of OECD and for the first time also for Spain, that there is a positive correlation between getting good academic results and certain non-cognitive skills, such as high levels of perseverance and ability to work or preference for complex problem solving. However, we can not guarantee the causal content of the estimated correlation and so it is likely that obtaining good academic results encourages investment, by the student and/or their family, in non-cognitive skills or personality traits that in turn lead to improved academic results.

Table 6.7. Determinants of the score in problem solving. OLS estimation.

	OECD		Spain	
Variable	Coefficient	t-Statistic	Coefficient	t-Statistic
Principal factor, student	9.68	25.21	10.61	8.15
Age	13.91	6.82	10.51	1.80
Woman	-10.19	8.80	-7.85	1.71
Higher studies (f)	5.85	1.85	7.82	1.13
Average studies (f)	6.96	2.25	21.55	3.46
Higher studies (m)	-7.21	1.94	-23.06	2.78
Average studies (m)	-5.24	1.61	-16.84	2.15
Occupation 1 (f)	10.83	4.79	10.86	1.66
Occupation 2 (f)	19.20	8.95	21.94	2.84
Occupation 3 (f)	12.03	4.16	6.99	0.75
Occupation 4 (f)	12.08	3.88	9.35	1.10
Occupation 5 (f)	3.26	1.34	-5.54	0.82
Occupation 6* (f)	-3.89	1.60	4.63	0.65
Occupation 8 (f)	-8.31	3.74	-1.94	0.22
Occupation 1 (m)	14.82	3.32	8.95	0.73
Occupation 2 (m)	18.76	6.48	15.34	1.62
Occupation 3 (m)	20.54	6.52	15.91	1.38
Occupation 4 (m)	18.46	5.43	11.45	1.18
Occupation 5 (m)	8.15	3.08	8.19	0.91
Occupation 6* (m)	14.33	3.11	-14.23	0.71
Occupation 8 (m)	-0.65	0.24	-7.11	0.77
Private school	-13.39	4.49	-1.90	0.14
Average city	5.03	1.28	6.13	0.51
Large city	1.25	0.38	0.49	0.04
Books: 11-25	16.71	7.25	24.10	2.74
Books: 26-100	34.00	15.16	51.08	6.01
Books: 101-200	46.45	18.99	79.38	9.20
Books: 201-500	58.12	23.06	88.23	9.51
Books: > 500	55.10	22.26	95.13	8.90
Language at home	2.18	0.59	7.35	1.04
Lack of teachers	-1.44	0.56	-7.68	0.54
Lack of resources	-1.42	0.35	-4.25	0.33
Autonomy hiring	-1.39	0.42	2.93	0.22
Autonomy salary	6.71	1.58	-11.32	0.73
Autonomy budget	1.63	0.50	10.81	1.10
Autonomy content	0.67	0.19	9.13	1.16
Qualified teachers	6.05	1.70	19.55	1.20
Schools nearby	-3.16	1.06	-26.63	2.02
Pupils by grade	-2.51	0.83	30.21	3.51
Evaluates teachers	-1.73	0.57	1.22	0.15
Immig. generation 2	2.56	0.82	-1.02	0.18
Immig. generation 1	-2.21	0.50	4.05	0.66
Preschool education	17.04	7.14	41.59	5.59
Socioec. index school	35.26	11.99	14.79	1.62
Constant	221.54	6.79	173.65	1.90
Goodness of fit	0.30	-	0.23	-
Observations	98135	-	5793	-

Note: (p) and (m) indicate that the variable in question refers to the father or mother of the respondent, respectively.

Having reached this point, and in order to take a step further in the analysis of the relationship between non-cognitive skills and educational performance, we turn to microdata from the World Values Survey (WVS). The WVS includes, from its first surveys in the mid-eighties of the last century, a question about the most relevant values or qualities to encourage in a child. This survey is conducted over a large number of countries, including those considered here, and asks the respondent, among other things, to select up to five of the following eleven personality traits to encourage in a child: good manners, independence, responsibility, hard work, imagination, tolerance and respect for others, austerity and ability to save, determination, perseverance, religious faith and generosity.

We use the first two waves of the WVS, conducted between the mid-eighties and early nineties of the last century, to identify the most valued non-cognitive skills in each OECD country at that time. To do this we calculate the proportion of people in each country that select each one of the eleven qualities as well as one of the five to encourage in a child and subsequently we get the first principal component of these proportions. Table 6.8 presents the corresponding weightings for the eleven personality traits included in the calculation of the first principal component.

Table 6.8. Relative importance of non-cognitive skills to instill in a child for obtaining the first principal component. PISA 2012. OECD countries

Qualities	Factor
Good manners	0.195
Independence	0.374
Hard work	-0.111
Responsibility	0.482
Imagination	0.242
Tolerance and respect	0.044
Saving ability	0.326
Perseverance	0.415
Religious faith	-0.248
Generosity	-0.282
Obedience	-0.313

The new variable provides a synthetic measurement of the most valued non-cognitive skills in each country. This measurement not only reflects the unconditioned preference of society of the country, but also the effect of the educational and work institutions, etc., and the economic conditions in place at the given times. In any case, the new variable provides potentially relevant information for our analysis and so it summarizes the most valuable non-cognitive skills in the country of residence of the students surveyed in PISA 2012, at least a decade before they were born. Thus, this new variable does not suffer from the potential endogeneity of the non-cognitive skills reported by the student.

The estimations in which we include the new variable in the set of regressors are presented in Table 6.9.[5] We find a significant positive correlation between the synthetic measurement of non-cognitive skills preferred in each country in the first two waves of the WVS and the grade obtained in problem solving by the student surveyed in PISA 2012. This result is particularly interesting and so the sample includes students not born in the country of residence and second-generation immigrants whose social standard with regard to non-cognitive skills to be encouraged in a child is probably different from the prevailing one in the destination country.

The results obtained in Tables 6.8 and 6.9 suggest that the countries that value responsibility, perseverance, independence, the ability to save and put off rewards, or imagination, as qualities to be encouraged in a child, are those that get better results in problem

[5]The results remain qualitatively unchanged even when we don't eliminate approximately a third of respondents from the initial sample of each country, from PISA 2012, who failed to answer the questions about perseverance and preference for problem solving.

solving. By contrast, countries that put an emphasis on obedience, generosity or religious faith get systematically worse results than those from the first group in educational performance.

In a related line of study, Méndez (2014) analyzes whether the grade obtained in PISA 2003, 2006, 2009 and 2012 for a sample of second generation immigrants is partly determined by the qualities considered essential to instill in a child in their parents' country of origin twenty years earlier, when it is more likely that parents were still living in their country of origin and, therefore, were influenced by those social preferences. These measurements were also obtained from the WVS.

Méndez (2014) finds that between a fifth and a quarter of the average differences in academic achievement in language, maths and sciences between the different countries of origin can be explained by the unequal provision of non-cognitive skills considered relevant to a child's upbringing. This result is obtained in all of the destination countries considered in that study, some of them characterized by very different educational systems. Also, this result shows that non-cognitive skills are transmitted across generations from parents to children, suggesting that investment in these skills or personality traits has higher rates of return than those estimated in conventional models that do not provide for this intergenerational transmission. In other words, the investment in these skills is, very probably, below the optimal level both from the private and social point of view.

Table 6.9. Non-cognitive skills and problem solving. MCO Estimation

Variable	Coefficient	t -Statistic
Principal factor, student	4.57	7.81
Age	13.96	6.56
Woman	-14.26	11.55
Higher studies (f)	10.88	3.48
Average studies (f)	9.13	2.98
Higher studies (m)	-5.46	1.55
Average studies (m)	-6.32	1.99
Occupation 1 (f)	15.54	6.66
Occupation 2 (f)	21.63	9.54
Occupation 3 (f)	16.63	5.55
Occupation 4 (f)	17.89	5.71
Occupation 5 (f)	5.24	2.10
Occupation 6* (f)	-2.69	1.12
Occupation 8 (f)	-4.32	1.88
Occupation 1 (m)	14.63	3.23
Occupation 2 (m)	17.86	5.89
Occupation 3 (m)	19.26	6.06
Occupation 4 (m)	19.17	5.53
Occupation 5 (m)	8.56	3.30
Occupation 6* (m)	13.88	2.91
Occupation 8 (m)	0.17	0.06
Private school	-9.29	3.20
Average city	5.56	1.36
Large city	6.88	2.00
Books: 11-25	17.62	7.48
Books: 26-100	38.96	16.97
Books: 101-200	52.97	19.59
Books: 201-500	68.31	27.05
Books: > 500	65.94	22.89
Language at home	4.85	1.27
Lack of teachers	-2.90	1.05
Lack of resources	-3.01	0.73
Autonomy hiring	-9.20	2.66
Autonomy salary	7.06	1.74
Autonomy budget	0.45	0.13
Autonomy content	3.57	1.03
Qualified teachers	10.39	2.95
Schools nearby	-3.97	1.23
Pupils by grade	-0.02	0.01
Evaluates teachers	3.28	1.06
Immig. generation 2	1.46	0.44
Immig. generation 1	-1.58	0.35
Preschool education	15.45	4.78
Socioec. index school	33.65	11.77
Constant	218.20	6.54
Goodness of fit	0.23	-
Observations	95897	-

Note: (p) and (m) indicate that the variable in question refers to the father or mother of the respondent, respectively. The number of observations is lower than in Table 6.7 because Israel did not participate in the first two waves of the WVS and therefore is excluded from this estimate.

Unlike Méndez (2014), in this study we can not isolate the mechanism of intergenerational transmission of non-cognitive skills not because we do not know the country of origin of the parents of the students surveyed in PISA. An analysis conditioned to the country of residence, in which the economic and institutional conditions were homogeneous for second generation immigrants living in that country, would allow us to check the hypothesis that cultural heritage in non-cognitive skills affects student achievement. That is precisely the analysis in Méndez (2014). Our synthetic measurement reflects the non-cognitive skills most valued in the country of residence of the student surveyed in PISA in a past moment in which most of the parents of the native students were living in the country and therefore, were influenced by that standard. However, our measurement of the evaluation of qualities or personality traits is also influenced by the educational and work institutions, etc., of each country and by the prevailing macroeconomic conditions in the past moment in which the survey was conducted. That is, our synthetic measurement of non-cognitive skills by country of residence also reflects the skills that enhance each specific national and historical combination of the education system, job market and economic conditions, not only the preferences of society.

It's worth highlighting that the synthetic measurement of non-cognitive skills obtained in Méndez (2014) assigns weightings, which are very similar to those obtained in this study and presented in Table 6.8, to each of the eleven qualities defined in the WVS. This result is interesting, especially when the countries considered in the two studies are very different.

The results obtained in this study can be placed in relation to the literature on personality psychology using the correspondence established in Méndez (2014) between qualities to encourage in a child and the taxonomy most frequently used to characterize the personality in the psychology literature, the "Big Five Personality Index". This classification system, resulting from the application of factorial analysis to a broad set of personality descriptors, distinguishes between five aspects or facets of individual personality: conscientiousness, agreeableness, extraversion, openness to experience and emotional instability. Méndez (2014) reviews the literature on the subject and concludes that responsibility, perseverance, the ability to postpone rewards, qualities with a positive effect on school performance, are closely related to the first factor. Meanwhile, imagination and independence are clearly related to the factors openness to experience and agreeableness, respectively. Thus, the results obtained in this study are in line with those previously obtained in the literature on the subject by pointing to the first factor, conscientiousness, as the most relevant in educational performance (Heckman, 2011).

Finally, Table 6.10 presents three new estimations in which we explore the relationship between non-cognitive skills and academic performance in the OECD in greater depth. In the first estimation we replace the synthetic measurement of non-cognitive skills obtained using the first two waves of the WVS by a similar measurement obtained using the following two waves, ie: we use surveys conducted in the early and mid-nineties. The idea behind this analysis is that if we are really picking up cultural and institutional aspects underlying societies analyzed with our synthetic measurement, the results should not change substantially in a decade, since culture is an institution characterized by slow change (Roland, 2010). The results from the synthetic measurement obtained from waves 3 and 4 of the WVS are virtually identical to those obtained using the two previous waves, something which can be interpreted as evidence in favor of the hypothesis that our synthetic measurement picks up social preferences conditioned by institutions and underlying situations in the subject of non-cognitive skills.

The second estimation includes the synthetic measurement of qualities to instill in a child obtained in the first two waves of WVS, and the synthetic measurement of perseverance self-reported by the student in PISA 2012, together as explanatory variables. The results suggest that the two measurements are relevant to student performance and are weakly correlated.

Finally, the third estimation uses the same specification as the previous estimation but analyzes the determinants of the maths grade instead of problem solving as a dependent variable. The results confirm that both the skills self-reported by the student as the most valued by the society of the country in which they live relevantly condition their academic performance in all subjects analyzed. In addition, the two synthetic measurements of non-cognitive skills are more related to the maths grade than to problem solving.

Table 6.10. Non-cognitive skills and problem solving. Tests of robustness

	Problem solving				Maths	
Variable	Coeff. (1)	t-stat.	Coeff. (2)	t-stat.	Coeff. (2)	t-stat.
Principal factor, country	4.64	7.47	5.79	9.49	8.25	13.00
Principal factor, student	-	-	7.26	19.52	9.66	26.64
Age	15.33	7.15	12.65	6.07	12.86	6.47
Woman	-14.26	11.18	-11.57	9.61	-13.32	10.41
Higher studies (f)	10.58	3.45	10.20	3.24	5.88	1.83
Average studies (f)	7.93	2.63	9.38	3.06	1.07	0.34
Higher studies (m)	-10.16	2.90	-3.64	0.98	-9.54	2.52
Average studies (m)	-10.34	3.26	-4.06	1.21	-8.07	2.53
Occupation 1 (f)	15.37	6.46	13.53	5.93	14.99	6.46
Occupation 2 (f)	22.29	9.49	20.33	9.28	24.62	12.66
Occupation 3 (f)	16.95	5.54	15.77	5.41	16.86	6.47
Occupation 4 (f)	20.38	6.54	18.01	5.87	21.66	6.95
Occupation 5 (f)	6.32	2.49	4.62	1.90	3.95	1.92
Occupation 6* (f)	-2.96	1.21	-2.90	1.20	-5.23	2.19
Occupation 8 (f)	-5.40	2.27	-3.57	1.60	-8.60	3.78
Occupation 1 (m)	14.21	3.00	12.32	2.80	11.51	3.00
Occupation 2 (m)	17.99	5.79	16.00	5.34	16.91	5.81
Occupation 3 (m)	19.88	6.02	17.17	5.52	15.93	5.72
Occupation 4 (m)	19.91	5.54	18.61	5.52	17.92	6.51
Occupation 5 (m)	9.11	3.38	8.22	3.21	5.36	2.03
Occupation 6* (m)	14.55	2.98	13.70	2.94	19.06	3.59
Occupation 8 (m)	1.77	0.65	-0.58	0.22	-3.27	1.21
Private school	-8.50	2.86	-7.55	2.59	-0.16	0.06
Average city	5.59	1.35	5.87	1.42	1.45	0.42
Large city	6.67	1.90	7.47	2.18	3.52	1.14
Books: 11-25	17.86	7.36	16.24	7.02	19.44	9.31
Books: 26-100	39.46	16.91	35.81	16.04	41.15	18.30
Books: 101-200	53.40	19.24	48.84	18.55	52.74	21.33
Books: 201-500	69.39	26.43	62.10	24.60	73.15	28.58
Books: > 500	66.73	22.17	58.49	21.55	72.66	28.14
Language at home	6.38	1.58	6.15	1.59	2.76	0.82
Lack of teachers	-3.81	1.33	-3.24	1.18	-5.56	2.42
Lack of resources	-2.60	0.61	-2.90	0.73	1.77	0.70
Autonomy hiring	-10.26	2.76	-11.57	3.37	-11.93	3.40
Autonomy salary	6.83	1.63	6.25	1.56	-5.24	1.69
Autonomy budget	-0.30	0.08	0.76	0.22	0.05	0.02
Autonomy content	2.57	0.70	5.28	1.54	8.15	2.53
Qualified teachers	10.94	3.03	10.16	2.80	1.20	0.41
Schools nearby	-3.85	1.16	-3.32	1.01	-0.73	0.27
Pupils by grade	-0.91	0.29	0.39	0.12	-8.73	2.41
Evaluates teachers	3.21	1.02	3.48	1.14	6.48	3.01
Immig. generation 2	-0.42	0.12	0.97	0.30	-0.82	0.26
Immig. generation 1	-4.14	0.85	-3.08	0.67	-5.44	1.36
Preschool education	9.00	2.78	18.67	6.05	14.73	4.86
Socioec. index school	31.38	10.35	34.04	11.66	38.05	13.73
Constant	206.03	6.08	236.27	7.17	252.48	7.85
Goodness of fit	0.24	-	0.25	-	0.34	-
Observations	86566	-	95897	-	95897	-

Note: (1) and (2) indicate that the most valued non-cognitive skills in a child were obtained from waves 3 and 4 of the World Values Survey or from the waves 1 and 2, respectively. Meanwhile, (p) and (m) indicate that the variable in question refers to the father or mother of the respondent, respectively. The difference in the number of observations between the first estimation and the other two is due to the fact that Canada did not participate in wave 3 and 4 of the WVS but Israel did.

Discussion

The above statistical and econometric analyses provide several fundamental results. The first of these is that Spain is below the OECD average in student performance, the adverse differential being higher in problem solving than that registered in a regulated subject like maths.

Furthermore, we find that Spain should, owing to its provision of student, family and school resource characteristics, register a higher relative position than that observed. In other words, the unfavorable relative position of Spain is not a problem of provision of resources, but that the organization of these is inefficient in Spain with respect to that observed in other developed countries. In this sense, a more detailed investigation is needed to point to which organizational aspects of the education system in Spain explain the lower performance of the provision of determinants in our country.

Another lesson we can draw from the analysis carried out in this chapter is that, as pointed out by recent scientific evidence at an international level, non-cognitive skills or personality traits are relevant determinants of student performance. This study is the first in verifying such relevance in Spain. Specifically, we have shown that there is a positive correlation between the score obtained in problem solving and the level of perseverance and preference for problem solving reported by students. This result should not be interpreted in a causal sense because of its obvious potential endogeneity, but it represents a first contribution from being able to determine, in future research, the meaning and causal content of the relationship between the two variables.

Another contribution which stands out, closely related to the previous one, is that we have shown that the social preferences reported regarding the qualities to encourage in a child several decades before the 2012 PISA exam have a relevant effect on student scores, even after controlling for a wide range of characteristics of the students, their family and their school and even after controlling for the level of non-cognitive skills reported by the student. The auxiliary analyses ruled out the existence of a positive correlation between the two measurements of non-cognitive skills, national and individual, something expected a priori if there really is intergenerational transfer of non-cognitive skills. However, there are several reasons why the result is not inconsistent with the fact that this transfer of non-cognitive skills between parents and children really occurs. Firstly, the cultural heritage in terms of non-cognitive skills of the students surveyed in PISA is determined by the country of birth of their parents, information that is not available to us. Therefore, we can not assign to each student the theoretical value from the synthetic measurement of non-cognitive skills obtained in the WVS sample of the country of origin of their father or mother and check the cultural hypothesis, ie: the intergenerational transmission of non-cognitive skills that has been shown previously in the literature in Méndez (2014) .

Secondly, the format of the questions is very different. So, while in PISA it only asks about perseverance and preference for problem solving, ie: without having to choose between those or other qualities, in the WVS respondents have to choose a maximum of five qualities of a list of eleven. It is therefore possible that a respondents in the WVS point to qualities other than perseverance as more relevant even though they report, if asked, high levels of perseverance or preference for problem solving.

Finally, the results obtained here are in line with those found in the literature on non-cognitive skills. Specifically, we find that students living in countries that in the eighties encouraged certain qualities in children such as responsibility, perseverance, independence, the ability to save and to put off rewards and imagination, through the family and their educational system, obtain better results both in maths and problem solving, the same as with personal, family and school characteristics. By contrast, students resident in countries that emphasize obedience, generosity or religious faith get systematically worse results in educational performance. Similar conclusions are obtained if we use the skills most valued in the students' countries of residence in the nineties. Since the measurements obtained from the WVS reflect the effect of economic conditions, the fact that the results do not depend on the decade in which those measurements are obtained reinforces their interpretation as variables that approximate the prevailing social standard in each country, the result of social and institutional preferences, in non-cognitive skills.

In order to provide a quantitative measurement of the importance of non-cognitive skills we can say that an increase equivalent to one standard deviation in the value of the national cultural synthetic measurement explains 7.5% and 14.5% of the dispersion in academic performance in problem solving and maths, respectively, among OECD countries. These numbers are relevant, especially if we consider that an equivalent increase in the proportion of parents employed in the most outstanding occupational category, the first, would explain 3.7% and 5.5%, respectively, of the observed differences in average score between OECD countries. The percentages obtained for a hypothetical increase in the proportion of parents with higher education in a measurement equivalent to a standard deviation are 3.5% and 2.7%, respectively.

Non-cognitive skills are therefore an efficient way of improving the education system. The next step must be to investigate how to modify the provision of non-cognitive skills existing in a country in order to redirect it towards those qualities that contribute to getting good results in education, career and health in adulthood.

This result provides a new explanation for the unfavorable relative position of Spain in terms of academic performance. In the first survey that the WVS conducted in Spain in 1981 the qualities highlighted by Spanish society as appropriate for a child were, in order of relative importance, the following: responsibility (63.1%), good manners (53.5%), tolerance and respect for others (44.2%), hard work (41.4%) and obedience (29.7%). Except for the first of the qualities, the most marked in that survey, all the others contribute marginally or negatively to the synthetic measurement of non-cognitive skills that has a positive effect on student achievement. By contrast, the most effective non-cognitive skills in increasing the performance of students such as perseverance, independence or the ability to save were only labeled as relevant by 12.6%, 24.3% and 10.7%, respectively, of the Spanish population surveyed in 1981. The figures do not change substantially if we use the survey from the early nineties of the last century, or even that performed in the middle of the first decade of this century.

The next question is: can we change the provision of non-cognitive skills of a society to adapt it to the optimal provision? The review of the literature suggests that we can, pointing to preschool interventions in children as being especially effective (Heckman, 2011; Almlund and others, 2011).

In this sense, the literature review conducted by the Department of Education of the United States (2013) on preschool interventions that promote key non-cognitive skills for educational success such as, for example, tenacity and perseverance, is particularly useful. The interventions are grouped into five categories: reading programs in schools that promote values; interventions that seek to change the mentality and strategies of student learning; alternative models of school; informal learning programs and digital learning programs which include tools for teachers. The common characteristic of the five categories is that this is a new way of learning and teaching, more individualized and more focused on meeting the student, working on their strengths, their non-cognitive skills, in order to obtain lasting or permanent results.

The next step for Spain is to learn from inside its education system, with partial initiatives, evaluated experimentally, which guarantee the success of future programs on a greater scale aimed at achieving the optimal qualities in Spanish students for their educational, career and health success in adulthood.

Conclusions

The aim of this chapter is to analyze the relative position of the countries of the OECD in problem solving in PISA 2012. To do this, we also study how much of the differences observed between countries are due to differences in provision of determinant factors of student performance. All of this by performing a distributional analysis which, unlike the conventional analyses focused on the average score of each country, uses information from the entire distribution of students by competence levels.

We find that Spain is below the OECD average in student performance, the unfavorable differential in problem solving being higher than that registered in a regulated subject like maths. This result is not explained by the provision of family and school factors of Spanish students, which are higher than the whole or average of the OECD. Thus, we find that Spain should register a higher relative position than that observed, because of its provision of student, family and school resource characteristics,.

The econometric analyses show that non-cognitive skills or personality traits are relevant determinants of student performance in problem solving and maths, especially in the latter competence. This study is the first in finding such relevance in Spain. Specifically, we have shown that there is a positive correlation between the score obtained in problem solving and the level of perseverance and preference for problem solving reported by students.

We also find that the social preferences regarding the qualities to encourage in a child reported several decades before the 2012 PISA exam have a significant effect on student scores, even after controlling for a broad set of student, family and school characteristics and even after controlling for the level of non-cognitive skills reported by the student.

Specifically, we found that students living in countries that encouraged certain qualities such as responsibility, perseverance, independence, the ability to save and to put off rewards and imagination in children, through the family and its educational system in the eighties, obtain the best results in both maths and problem solving, for the same personal, family and school characteristics.

By contrast, students resident in countries that emphasize obedience, generosity or religious faith, get systematically worse results in educational performance. Similar conclusions are reached if we use the skills most valued in the countries of residence of students in the nineties.

This result provides a new explanation for the unfavorable relative position of Spain in academic performance, since the personality traits most valued by the Spanish society are not those that increase the likelihood of educational, career and health success in adulthood.

REFERENCES

ALMLUND, M., DUCKWORTH, A., HECKMAN, J.J., y KAUTZ, T.(2011). Personality Psychology and Economics. En *Handbook of the Economics of Education*. Vol. 4, ed. E. A. Hanushek, S. Machin y L. WöBmann, 1-181. Amsterdam:Elsevier.

CONTI, G., HECKMAN, J.J., y URZÚA S. (2011). "Early Endowments, Education and Health. Human Capital and Economic Opportunity: A Global Working Group". Working Paper 2011-001.

DEPARTAMENTO DE EDUCACIÓN DE ESTADOS UNIDOS (2013). Promoting Grit, Tenacity, and Perseverance: Critical Factors for Success in the 21st Century. Office of Educational Technology.

FELFE, C., NOLLENBERGER, N. y RODRÍGUEZ-PLANAS, N. (2012). Can't buy my Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development", *IZA Discussion Paper* 7053.

HANUSHEK E.A., y WOESSMANN L. (2011). The Economics of International Differences. En Educational Achievement. *Handbook of the Economics of Education*, Elsevier.

HECKMAN, J.J. (2008). "Schools, Skills and Synapses". Economic Inquiry 46 (3), pp. 289-324.

HECKMAN, J.J. (2011). "Integrating Personality Psicology into Economics". *NBER Working Papers* 17378.

HECKMAN, J.J., R. PINTO y SAVELYEV, P. (2012). "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6), pp. 2052-86.

HERRERO, C., MÉNDEZ, I., y VILLAR, A.(2014). "Analysis of group performance with categorical data when agents are heterogeneous: The evaluation of scholastic performance in the OECD through PISA". *Economics of Education Review*. doi:http://www.sciencedirect.com/science/article/pii/S0272775714000193

HERRERO, C., y VILLAR, A. (2013). "On the Comparison of Group Performance with Categorical Data". *PLoS ONE* 8(12): e84784. doi:10.1371/journal.pone.0084784

MARCERANO, O. (2014) "Del lápiz al ordenador, ¿diferentes formas de evaluar las competencias del alumnado?". En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MEDIAVILLA, M., y ESCARDÍVUL, J.O. (2014) "Efectos de las TICs en la adquisición de competenciasUn análisis de género y titularidad de centro para las evaluaciones por ordenador". En INEE (ed.): PISA 2012: Resolución de problemas. Informe Español. Volumen II: Análisis secundario, Madrid: Instituto Nacional de Evaluación Educativa.

MÉNDEZ, I. (2014). "Culture, noncognitive skills and student performance". Mimeo.

OECD (2009). PISA 2006 Technical Report. Paris, OECD.

OECD (2011). PISA in Focus 2011/1.

ROLAND, G. (2008). "Understanding institutional change: Fast-moving and slow-moving institutions" *Studies in Comparative International Development 38 (4)*, pp. 109-131.

PISA 2012

GOBIERNO DE ESPAÑA
MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE