



Facultad  
de Ciencias  
Económicas y  
Empresariales

Departamento  
de Economía  
Aplicada y  
Estadística



**Revista de Evaluación de  
Programas y Políticas Públicas**  
JOURNAL OF PUBLIC PROGRAMS AND POLICY EVALUATION

# The Determinants of success in Primary Education in Spain

Núm. 2 (2014), pp. 22-53.

Brindusa Anghel<sup>1</sup>

Antonio Cabrales<sup>2</sup>

Recibido: Enero, **2014**

Aceptado: Marzo, **2014**

**JEL Clasif:** JEL-class: I20, I21, I28

---

<sup>1</sup> Brindusa Anghel would like to gratefully acknowledge the financial help by the Ministry of Science and Innovation, Spain, under the grant ECO2009-07530. FEDEA [banghel@fedea.es](mailto:banghel@fedea.es)

<sup>2</sup> Universidad Carlos III de Madrid and FEDEA. Department of Economics. Universidad Carlos III de Madrid. Madrid. Spain. [antonio.cabrales@uc3m.es](mailto:antonio.cabrales@uc3m.es)

### **Abstract**

This paper studies the importance of school inputs, social environment and parental background on the outcomes of a standardized exam for sixth grade students in the region of Madrid, Spain. These data can then be linked to individual characteristics such as nationality, family type and parental education and profession. For public schools we also have school level data on variables such as number of students, teachers and some of their characteristics, students' nationalities, extra-curricular activities, and the number of students needing special assistance

Key Words: Education, Human capital, Standardized tests

### **Resumen**

Este trabajo analiza la importancia de los recursos de los colegios, del nivel socioeconómico y de las características de los padres sobre los resultados de un examen estandarizado para los alumnos del 6o curso de educación primaria de la Comunidad de Madrid. Estos datos se pueden unir con características individuales como la nacionalidad, la composición de la familia y la educación y la ocupación de los padres. Asimismo, para los colegios públicos tenemos datos a nivel de colegio de variables como el número de alumnos, profesores y algunas de sus características, la nacionalidad de los alumnos, actividades extraescolares y el número de alumnos que requieren atención especial.

Palabras clave: Educación, Capital humano, Exámenes estandarizados.

## 1. Introduction

The productivity of an economy depends in a crucial way on the quantity and quality of the human capital accumulated by the labor force which that economy can access. Although human capital can be in part “imported” by allowing sizable immigration flows, as Spain did in the last decade, most of it will necessarily be “home-grown”. Thus, the education system, is an essential determinant of the productivity of any economy. Furthermore, education is also important to achieve the right allocation of talent to the productive system, and hence of social mobility.

The aim of this paper is to contribute to the debate on education in Spain. After a brief review of the international evidence, we analyze a database arising from a standardized exam that the Comunidad Autónoma de Madrid (Madrid regional government) has been conducting since 2005 for all 6th grade students in the region, who are hence in the final year of primary school. We choose to analyze the region of Madrid because of the availability of this database for student achievement in primary education in Spain. Moreover, the region of Madrid is one of the most important regions in Spain: in 2011 in terms of income per capita and in terms of the population it is the third region in Spain.

The exam measures what the authorities consider basic skills (Competencias y Destrezas Indispensables, or C.D.I.) in mathematics, language, dictation, reading and general knowledge. The exam does not have academic consequences for the children. Like the tests conducted by the PISA program, the C.D.I. exam serves mostly an informative purpose, both for the student and for the authorities. In spite of this, it is possible that the exam conditions the teaching strategies and outcomes, since its results are made public and the teachers and principals could feel accountable for them.

From the exam we have information on a few individual characteristics of the children, like gender, nationality, whether the student has any disability and whether the student has any special educational needs. In the year 2008/09, students were asked to fill in a short questionnaire from which we have information on additional individual-level controls, like parents’ level of education and profession. Therefore, we will focus our empirical analysis on the data for the 2008/09 cohort.

With respect to our study of the Madrid data, there are several important findings to be noted. First, we summarize briefly the effects of personal characteristics, which are available for the 2008/09 data. Parental education and profession are extremely important, confirming results from earlier studies. Among these two factors, education is clearly more important than profession, but the two have significant independent impact. Children living with single parents do worse than the rest, and having brothers and sisters tends to be associated with good performance.

Immigrants sometimes do worse than nationals, even after controlling for parental education, but this is far from universal. Some of this is expected, as Asians or Eastern Europeans tend to do very well in mathematics and less well in language (in fact, poorly, in the case of Asians). But some of this is surprising. Moroccan students are no different from Spaniards in mathematics and general knowledge, once parental background is controlled for, although a bit worse in language and dictation. Latin Americans are the large group with the worst outcomes, even after taking into account parental background. They do worse than Spaniards in all subjects. For example, their Spanish scores are at the level of those for East Asians.

In terms of school level variables the first conclusion is that class size appears to have no significant impact on students' achievement. This is important because the usual response by some sectors to educational problems is: more resources are needed. The evidence on class size implies that extra resources can be easily misspent on unproductive policies.

Another useful result is that the percentage of immigrants in the class has at best a small effect on the performance of other children, and it is not statistically significant in most of the regressions. At worst, we find that in dictation schools that have between 21% and 30% immigrants in the 6th grade perform slightly better than schools that have more than 40% immigrants in the 6th grade.

We also find that the average parental education in the class has a statistically significant impact on the school fixed effect. That is, beyond the effect of their own parental background, the children experience an additional effect coming from their peers' parental background. This could explain why the unconditional means of schools with large amounts of immigrants are lower, even though the proportion of immigrants

has no explanatory power for the fixed effect. Those badly performing schools with lots of immigrants also concentrate large amounts of children with less educated parents.

A noteworthy aspect of this project is that we have found some concrete evidence of the importance on parental involvement in the school. The percentage of school activities organized by the Parents' Association (AMPA) has a large and significant effect on the dictation, language and mathematics exams.

The school type (public, private or "concertado") is important before we control for the family background of class members. More precisely, the impact of private and "concertado" status on the school fixed effect is highly significant. However, once we control for the socioeconomic composition of the classroom the effect vanishes. This means that these types of schools have been good at separating the children of better educated parents from the rest, and allowing their students to extract the "peer effects" of their classmates.

The remainder of the paper is organized as follows. In section 2 we discuss the relevant literature on the economics of education. Section 3 describes the data and the empirical strategy. Section 4 shows the results. Section 5 concludes.

## 2. The Literature

The literature on the determinants of students's achievement is extensive. We can distinguish between studies that examine the relationship between individual characteristics and/or family background and educational achievement and studies on school characteristics and students' academic performance. Both of these lines of research have received large attention in the literature on economics of education.

Ermisch and Francesconi (2001) find that parents' educational attainments are strongly associated with their children's educational attainments. They also find that, in addition, young adults who experience a single parenthood as children and those who come from families in the bottom income quartile have significantly lower educational attainments. Woessmann (2004) using data from the Third International Mathematics and Science Study (TIMSS) estimates the impact of parental education and other measures of family background on children's test scores. They find that the education

level attained by the parents is strongly related to student performance in all European countries, as well as in the United States. Fertig (2003), in a study of the determinants of German Students' Achievement in the PISA 2000 Study, finds that many individual and family background characteristics have a statistically significant impact on individual test scores.

The literature on policy interventions in the educational sector is also vast. School resources is probably the kind of intervention that has received the largest amount of attention by economists. There are a couple of reasons for this. One is that it is easy to measure with the current databases. The money spent per student, the number of teachers in a school, are typically available in the data which economists have been able to use. Another important reason is that a first line of defence for administrators whose schools are seen to be floundering is the rallying cry of: "we need more resources".

Unfortunately, the evidence on this is rather mixed. On a first approximation the case for adding resources or reducing class sizes to improve school performance does not look very strong. Class sizes in many countries have been reduced, dramatically in some cases, over the last few decades, without any noticeable improvement in student performance, measured through standardized testing. For example, Hanushek (1999, 2003) reviews the literature and fails to observe consistent improvements in quality through input-based policies.

This conclusion is controversial. Much of the evidence, as we mentioned, comes from longitudinal databases where it is often difficult to be sure that all relevant variables are included, and firmly establishing causality can be very difficult. For this reason, some researchers have resorted to other forms of analysis that can determine the causal effect of class size in a more robust way.

Probably the least controversial method for analyzing the evidence on class size, or any other policy, is through a field experiment. A group of students is randomly selected to be "treated" with a smaller class size and their performance is compared to the one of a control group. The most famous experiment in this field is the Tennessee STAR experiment. Finn and Achilles (1990) report gains of about 0.3 to 0.6 of a

standard deviation in the class mean. Krueger (1999) reanalyzes the data with more sophisticated econometric techniques and finds somewhat larger gains.<sup>3</sup>

An alternative way of analyzing the problem that has the advantage of being more widely applicable was introduced by Angrist and Lavy (1999). The authors use the fact that in Israeli schools when the class reaches a ceiling of 40 students, it must be split into two smaller groups. One can use this exogenous split to create instrumental variables, and also to produce a regression discontinuity design, where outcomes are compared at both sides of this boundary. They find an effect that is sizeable. The grades increase by about a third of a standard deviation in the class mean, for reducing the number of students by eight. The effect is clear among fifth graders and (a little bit less) for fourth graders, but virtually non-existent for third graders.

Other researchers, with data from different countries, such as Hoxby (2000) for the U.S. and Leuven, Oosterbeek, Rønning (2008) (and even this paper, for Spain) show that, for their datasets, the effect of class size is statistically indistinguishable from zero. Given the larger initial class size in Israel than in the other places it is not difficult to envisage a situation where class sizes are important at a high level, but become less significant when the base level is reduced.

From this short review it is easy to see that the issue remains contentious, and it is not completely settled. Under these circumstances it is difficult to make a recommendation in favour of expensive policies for class size reduction. In the case of Spain, the main policy tool to deal with students with particularly bad backgrounds or circumstances (“Estudiantes con necesidades educativas especiales” or “Educación compensatoria”) is precisely to teach these students in radically reduced groups (often down to two or three students) for a large fraction of the day. Of course, this is quite different from reducing by a small amount the class size for a more standard student. Nevertheless, given their cost these programs need to be carefully evaluated.

---

<sup>3</sup> Hanushek (1999b) shows, however, that the experiment was not completely clean. There was attrition from the program (some students “stopped” the treatment) and significant crossover between classes. Furthermore, teachers and students knew they were participating in an experiment, and this could have affected the results.

### 3. Institutional Background and Description of the Data

The Spanish public educational system is composed of 6 years of primary school, 4 years of compulsory secondary education (E.S.O.) and 2 years of non-compulsory education, which is divided into vocational training (*ciclos formativos*) and preparation for college (*bachillerato*). There are also three years of free publicly funded pre-school, from ages 3 to 5. The pre-school children share the premises with those in primary school. Also, the pre-schoolers in one location have precedence over other children applying to the same primary school. As a consequence of this precedence rule most students at the primary level come from the preschool in the same location. The parents of children submit a list of school choices to the authorities. If there is no excess of demand for the school they have chosen, they are admitted. If there is excess of demand, the admission is based on criteria like proximity of the family home to the school and family income.

The data for our empirical analysis comes from a standardized exam that is administered each year to 6<sup>th</sup> grade students<sup>4</sup> (around 12-13 years old<sup>5</sup>) of all primary schools (about 1200) in the region of Madrid. The exam is called “Indispensable Knowledge and Skills exam” (*prueba de Conocimientos y Destrezas Indispensables*). This exam was introduced by the Education Council of the region of Madrid in the academic year 2004/05, and it is compulsory for all primary schools (public, private or charter). The exam does not have academic consequences for the children. It is intended to give additional information to the children and their families as well as to the educational authorities.

The exam consists of two parts of 45 minutes each: the first part includes tests of dictation, reading, language and general knowledge and the second part is composed of

---

<sup>4</sup> The Spanish educational system is composed of 6 years of primary school, 4 years of compulsory secondary education (E.S.O.) and 2 years of non-compulsory education, which is divided into vocational training (*ciclos formativos*) and preparation for college (*bachillerato*).

<sup>5</sup> In primary school, students can repeat a grade in case their performance is deemed insufficient. On average, in the whole of Spain, the percentage of repeaters is 6.2%. Madrid is close to the national average with 6.5%. For more statistics and details on the Spanish educational system, see e.g.

<http://www.institutodeevaluacion.mec.es/contenidos/indicadores/ind2009.pdf>.



mathematics exercises. Our measures of student achievement are the standardized scores to the yearly mean in each of these five subjects.

Additionally, in 2008/09 a short questionnaire was filled out by each student. In the questionnaire the students were asked a few questions about themselves, their parents and the environment in which they are living. The answers to this questionnaire provide rich information on individual characteristics of students that is not available for previous cohorts; therefore we focus our empirical analysis on this cohort.

Since the performance of children is a combination of individual and family characteristics and school resources, we distinguish two categories of control variables: variables at individual level (student characteristics and family background) and variables at school level. The availability of the data at these two levels of aggregation will justify our econometric methodology (see next section).

The individual level variables that can be obtained from the exam are: gender, nationality (Spanish or immigrant), whether the student has special educational needs and whether the student has any disability.

For the 2008/09 cohort the set of control variables at individual level is significantly larger. The questionnaire provides the following variables: the age of the student, the country of birth (Spain, China, Latin America, Morocco, Romania and other), the level of education of the parents, the occupation of the parents, the composition of the household in which the student lives and the age at which the student started to go to the school.

Regarding the education of the parents, students were asked to provide this information for both the mother and the father. However, to facilitate the interpretation we choose the highest level of education between the mother and the father.<sup>6</sup> We distinguish the following categories: university education, higher secondary education, vocational training, lower secondary education and unfinished compulsory education.

---

<sup>6</sup> Results are robust to using only the education of the mother, or separately that of the father and the mother.

In the case of the occupation of the parents, we apply the same strategy as for the education: we choose the highest level of occupation between the mother and the father. Accordingly, we differentiate between the following categories: professional occupations (teacher, researcher, doctor, engineer, lawyer, psychologist, artist, etc.), business and administrative occupations (CEO, civil servants, etc.) and low-skilled occupations (shop-assistant, fireman, construction worker, cleaning staff, etc.).

The composition of the household in which the student lives is constructed according to the answers to the question: “With whom do you usually live?”. We differentiate the following 7 categories: lives only with the mother, lives with the mother and one brother/sister, lives with the mother and more than one brother/sister, lives with the mother and the father, lives with the mother and the father and one brother/sister, lives with the mother and the father and more than one brother/sister and other situations.

The school level variables that are available for all schools (public, private and charter) are the following: class size, enrollment in the 6th grade and the geographical location of the school in the region of Madrid (East, West, North, South or Capital). Additionally, from the data at student level we compute the share of immigrants in the class and the shares of students with parents with a certain level of education in the class (university, higher secondary, vocational training, lower secondary and didn't finish compulsory studies).

The school level variables that are available only for public schools are the following: teacher/student ratio, average age of the teachers, share of extra-curricular activities organized by Parents' Associations, share of students that are eligible for a free meal, school size, share of students with special educational needs.

Class size is calculated as the total number of students enrolled in the 6th grade divided by the total number of 6th grade classes. Enrollment in the 6th grade is the total number of students registered in the 6th grade. Teacher/student ratio is the ratio of the number of professors teaching in the 6th grade classes and the number of students enrolled in the 6th grade. Average age of the teachers is the average age of the professors teaching in the 6th grade. Share of extracurricular activities organized by Parents' Associations is the percentage of these activities in the total number of extracurricular activities in a school. Share of students that qualify for a free meal and share of students

with special educational needs is the percentage of students in the 6<sup>th</sup> grade with one of these characteristics.

The dataset for the 2008/09 cohort is formed of 56,929 students in 1,227 public, private and charter schools. Out of this number of schools, 735 are public schools. However, due to data availability, we use in our estimations a sample of about 44,500 students in 1,222 schools for the individual level regressions and a sample of 558 public schools for the school level estimations.<sup>7</sup>

The descriptive statistics of the student and school data are presented in Tables 1 and 2.

**Table 1**  
**Descriptive statistics of individual level variables**  
**2008/09 cohort**

Variable	Mean	Std. Dev.	Min	Max
<i>Subjects</i>				
Dictation	0,12	0,94	-1,60	1,27
Mathematics	0,13	0,94	-1,77	1,94
Language	0,15	0,90	-2,04	1,71
Reading	0,14	0,91	-1,97	1,39
General knowledge	0,12	0,94	-1,66	1,88
<i>Individual characteristics</i>				
Female	0,49	0,50	0	1
Student with special educational needs	0,06	0,23	0	1
Student with disability	0,02	0,14	0	1
Student's age	12,13	0,40	10	17
Student Spain	0,82	0,38	0	1
Student Romania	0,02	0,15	0	1
Student Morroco	0,01	0,09	0	1
Student Latin America	0,10	0,30	0	1
Student China	0,00	0,07	0	1
Student other	0,04	0,20	0	1
Parent education - Univ.	0,48	0,50	0	1
Parent education - Higher secondary	0,17	0,38	0	1
Parent education - Vocational training	0,12	0,32	0	1
Parent education - Lower secondary	0,17	0,38	0	1
Parent education - didn't finish compulsory studies	0,06	0,23	0	1

<sup>7</sup>The data set for the 2005/06-2007/08 cohorts is composed of a total of 155,226 students in 1,237 public, private and charter schools (around 50,000 students in each year). Out of these schools, 735 are public schools.

*The Determinants of success in Primary Education in Spain*

Parent occupation - Business, minister, city hall, CCAA	0,22	0,42	0	1
Parent occupation- Professional	0,34	0,47	0	1
Parent occupation - Blue Collar	0,44	0,50	0	1
Lives only with the mother	0,07	0,25	0	1
Lives with the mother and one brother/sister	0,04	0,20	0	1
Lives with the mother and more than one brother/sister	0,02	0,13	0	1
Lives with the mother and the father	0,16	0,37	0	1
Lives with the mother and the father and one brother/sister	0,43	0,50	0	1
Lives with the mother and the father and more than one brother/sister	0,17	0,37	0	1
Other situations	0,11	0,32	0	1
Start school before 3	0,53	0,50	0	1
Kindergarden between 3 and 5	0,44	0,50	0	1
Start school at 6	0,02	0,15	0	1
Start school at 7 or more	0,01	0,11	0	1
Observations	44542			

**Table 2**  
**Descriptive statistics of school level variables 2008/09 cohort**

Variable	Mean	Std. Dev.	Min	Max
<i>Subjects</i>				
Fixed effects - Dictation	0.05	0.43	-1.50	1.43
Fixed effects - Mathematics	-0.03	0.36	-1.60	1.62
Fixed effects - Language	-0.07	0.38	-1.56	0.99
Fixed effects - Reading	-0.16	0.35	-1.43	1.11
Fixed effects - General knowledge	-0.26	0.45	-1.41	1.72
<i>School characteristics</i>				
<i>All schools</i>				
Class size	23.40	3.89	2.00	34.00
Enrollment 6th grade	49.49	23.37	2	177
% Immigrant students in 6th grade (0-10%)	0.38	0.48	0	1
% Immigrant students in 6th grade (11-20%)	0.28	0.45	0	1
% Immigrant students in 6th grade (21-30%)	0.17	0.37	0	1
% Immigrant students in 6th grade (31-40%)	0.09	0.28	0	1
% Immigrant students in 6th grade (>40%)	0.09	0.28	0	1
% Parent's education - Univ.	34.25	20.02	0.00	100.00
% Parents' education - Higher secondary	15.70	8.19	0.00	52.38
% Parents' education - Vocational training	10.02	6.17	0.00	62.50
% Parents' education -Lower secondary	16.16	10.78	0.00	75.00
% Parents' education - didn't finish compulsory studies	5.52	5.87	0.00	72.00
School Capital	0.40	0.49	0	1
School East	0.14	0.35	0	1
School North	0.08	0.26	0	1
School West	0.10	0.30	0	1
School South	0.29	0.45	0	1
<i>Public schools</i>				
Teacher/student ratio	0.26	0.15	0.08	1.67
Av. age teacher	43.09	4.46	29.11	57.78
No. AMPA activ./total no. of activ.	0.06	0.15	0.00	1.00
% Students free meal	13.23	9.28	0.15	48.57
School size	382.71	149.34	18.00	1088.00
% Students with special educational needs	11.40	10.63	0.00	55.56

## 4. Econometric Methodology and Results

Since we handle data at two levels of aggregation (data at individual level and data at school level), we use a two-stage estimation procedure in order to model the relationship between student and school characteristics and academic outcomes, a strategy that has been often followed by previous research (Loeb and Bound, 1996, Hanushek et al., 1996).

The first stage is an OLS regression in which the standardized grades of the students in all the schools are regressed against individual characteristics and school dummies (school fixed effects). The coefficients on the school dummies can be interpreted as the value added of the school, once differences in student characteristics are controlled for. The equation that we estimate is the following:

$$Y_{ij} = \alpha + \beta X_{ij} + \sum_j \delta_j D_{ij} + u_{ij}$$

where  $Y$  is the score of student  $i$  in school  $j$ .  $X_{ij}$  are individual characteristics of the student described in the previous section (like gender, nationality, etc.).  $D_{ij}$  is a school dummy variable that equals 1 if student  $i$  attends school  $j$  in the academic year 2008/09.

The second stage is an OLS regression in which the coefficients of the school dummies ( $\delta_j$ ) are regressed against the school level variables:

$$\delta_j = \gamma + \theta Z_j + v_j$$

where  $Z_j$  are school level variables. Because of large differences among schools sizes, we believe that there could be potential efficiency gains from weighting the data by an estimate of the covariance matrix. Therefore, the second stage estimations are weighted by the inverse of the estimated variance of the school fixed effects from the first stage.<sup>8</sup>

---

<sup>8</sup> We also ran the unweighted regressions, however results did not change significantly. They are available upon request.

#### 4.1. First stage results

Table 3 reports the results for the first stage regressions, for the 2008/09 cohort. The dependent variables are the individual standardized grades in each of the five tests. All regressions include school dummy variables  $D_{ij}$ . Their coefficients  $\delta_j$  are the average school effects once we account for individual characteristics.

**Table 3**  
Pooled OLS with school fixed effects (1st stage) for 2008/09

	Dictation	Mathematics	Reading	General knowledge	Language
	(1)	(2)	(3)	(4)	(5)
Female	0.213*** (0.008)	-0.148*** (0.008)	-0.003 (0.008)	-0.181*** (0.009)	0.090*** (0.007)
Student with special educational needs	-0.761*** (0.021)	-0.736*** (0.020)	-0.561*** (0.021)	-0.551*** (0.020)	-0.799*** (0.019)
Student with disability	-0.887*** (0.031)	-1.011*** (0.030)	-0.877*** (0.033)	-0.758*** (0.032)	-1.024*** (0.031)
Student's age	-0.330*** (0.013)	-0.314*** (0.013)	-0.211*** (0.012)	-0.245*** (0.012)	-0.339*** (0.012)
Student Romania	0.079** (0.032)	0.081*** (0.030)	0.053* (0.031)	0.045 (0.031)	0.078*** (0.030)
Student Morroco	-0.194*** (0.042)	0.014 (0.039)	-0.153*** (0.045)	-0.030 (0.046)	-0.177*** (0.039)
Student Latin America	-0.273*** (0.016)	-0.199*** (0.015)	-0.064*** (0.015)	-0.206*** (0.016)	-0.251*** (0.014)
Student China	-0.228*** (0.066)	0.500*** (0.077)	-0.121 (0.075)	-0.198*** (0.068)	-0.235*** (0.068)
Student other	-0.077*** (0.020)	-0.106*** (0.021)	-0.013 (0.020)	-0.097*** (0.020)	-0.080*** (0.018)
Parent education - Univ.	0.160*** (0.021)	0.261*** (0.019)	0.194*** (0.022)	0.214*** (0.019)	0.215*** (0.020)
Parent education - Higher secondary	0.099*** (0.021)	0.148*** (0.018)	0.131*** (0.022)	0.153*** (0.019)	0.142*** (0.019)

**Table 3. Pooled OLS with school fixed effects (1st stage) for 2008/09 (Cont.)**

	Dictation	Mathematics	Reading	General knowledge	Language
	(1)	(2)	(3)	(4)	(5)
Parent education - Vocational training	0.083*** (0.023)	0.135*** (0.020)	0.142*** (0.023)	0.159*** (0.021)	0.136*** (0.021)
Parent education - Lower secondary	0.051** (0.021)	0.080*** (0.018)	0.078*** (0.022)	0.072*** (0.019)	0.074*** (0.019)
Parent occupation - Business, minister, city hall, CCAA	0.075*** (0.012)	0.090*** (0.012)	0.065*** (0.011)	0.093*** (0.012)	0.092*** (0.010)
Parent occupation- Professional	0.132*** (0.011)	0.163*** (0.011)	0.122*** (0.011)	0.157*** (0.012)	0.161*** (0.010)
Lives only with the mother	-0.042 (0.031)	-0.058* (0.030)	-0.032 (0.032)	-0.054 (0.033)	-0.050* (0.028)
Lives with the mother and one brother/sister	0.025 (0.031)	0.079** (0.032)	0.062* (0.034)	0.030 (0.035)	0.042 (0.029)
Lives with the mother and the father	0.081*** (0.029)	0.068** (0.028)	0.047 (0.030)	0.091*** (0.032)	0.090*** (0.026)
Lives with the mother and the father and one brother/sister	0.137*** (0.028)	0.168*** (0.027)	0.097*** (0.029)	0.119*** (0.031)	0.148*** (0.025)
Lives with the mother and the father and more than one brother/sister	0.058** (0.029)	0.143*** (0.028)	0.098*** (0.030)	0.078** (0.031)	0.086*** (0.027)
Other situations	0.025 (0.029)	0.076*** (0.029)	0.032 (0.031)	0.039 (0.032)	0.035 (0.026)
Kindergarden between 3 and 5	-0.005 (0.008)	-0.057*** (0.008)	-0.037*** (0.008)	-0.047*** (0.009)	-0.026*** (0.007)
Start school at 6	-0.102*** (0.026)	-0.209*** (0.026)	-0.168*** (0.028)	-0.176*** (0.026)	-0.160*** (0.024)
Start school at 7 or more	-0.150*** (0.037)	-0.250*** (0.035)	-0.180*** (0.037)	-0.159*** (0.039)	-0.190*** (0.036)
Constant	3.865*** (0.161)	3.750*** (0.163)	2.507*** (0.151)	2.963*** (0.153)	4.008*** (0.154)
Observations	44542	44518	44542	44542	44542
R-squared	0.404	0.351	0.271	0.317	0.434

Notes:

1. Standard errors clustered at school level in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%
2. Base categories for dummies: male, student Spain, parent education - didn't finish compulsory studies, parent occupation - blue-collar, lives with the mother and more than one sister/brother, kindergarden less than 3



Results are in general robust among the five tests, however some differences arise.

In particular, we find that girls do better in dictation and language than boys: on average, girls get about 0.2 standard deviations more in dictation and 0.09 standard deviations more in language than boys. On the other hand, in mathematics, boys get on average 0.15 standard deviations more than girls, and in general knowledge they get 0.18 standard deviations more than girls.<sup>9</sup> Children with special educational needs and children with any disability perform, on average, quite poorly: they get between a half and one standard deviation less in all the subjects.

Since we can control for the country of birth of the immigrant children, we could identify differences by the country of birth of the children. We distinguish the following groups: Latin America (which includes Ecuador, Colombia, Peru, Bolivia, and Dominican Republic), Romania, Morocco, China and other countries. The groups that we choose correspond to the largest groups of immigrants in Spain, and, also, from the region of Madrid. According to the census in Spain at January 2009, out of the 5.6 million foreigners in Spain around 1 million lived in the region of Madrid. From them, around 33% were from Latin America countries, 19% were from Romania, 8% were from Morocco and around 4% were from China (these groups accounted for approximately 63% of the total of immigrants in the region of Madrid).<sup>10</sup>

We find that once we control for individual characteristics like parental education and occupation, on average immigrants do worse than nationals, however this is not always the case. Children from China do very well in mathematics (approximately one half of a standard deviation better than Spanish children), while children from Romania perform relatively better in dictation, mathematics and language (around 0.08 standard deviations better than Spanish children).

On the other hand, we find no significant difference between Moroccan students and Spanish students in mathematics and general knowledge, while in dictation, reading and language Moroccan perform worse. Children from Latin America are the group with the worst outcomes. They do significantly worse than Spaniards in all the

---

<sup>9</sup> This result is not so surprising if we look at the international results of TIMSS 2007 (Trends in International Mathematics and Science Study) in mathematics. In the 4<sup>th</sup> grade, boys had higher average achievement than girls in 12 countries, including United States, Sweden, Norway, Scotland, Netherlands, Germany, Austria, and Italy.

<sup>10</sup> These numbers take into account the nationality of the person.

parts of the exam. More concretely, the difference between Latin American and Spanish students equals around 0.25-0.27 standard deviations less in dictation and language and around 0.2 standard deviations less in mathematics and general knowledge.

The age of the student appears to matter as well: the younger the child, the better he performs in all the tests. The effect of the age is negative and statistically significant even after controlling for nationality (there might be immigrants that do not have the necessary knowledge to be in a grade according to their age) and for children with special educational needs or children with some disability. Naturally this is probably induced by students who needed to repeat a particular grade because of their poor performance.

Both parental education and occupation are extremely important, confirming results from previous studies. If we compare the magnitude of the coefficients, the level of education appears to matter more than the profession of the parents. The effect of education is the highest in the case of children whose parents have a university degree, in all the subjects: in mathematics, the difference between a child with a parent that has a university degree and a child with a parent that has no studies is about 0.2 standard deviations. This difference decreases in magnitude as the level of education of the parents decreases, but it remains statistically significant.

Students whose parents have professional occupations (teacher, researcher, doctor, engineer, lawyer, psychologist, artist, etc.) do significantly better than the rest. The coefficient for this dummy is approximately twice the coefficient for the dummy for parents with white collar occupations (like CEOs or civil servants). This might indicate that parents with professional occupations are likely to place a greater value on education than the rest.

The estimations reveal also that living with both parents is associated with better performance of children in schools. Moreover, it appears that having brothers or sisters is also related to improved student outcomes.

There is also empirical evidence for the fact that starting to go to school at an early age is beneficial for the school performance of children. A child that has started to go to school at less than 3 years of age gets around 0.2 standard deviations more in the CDI exam than a child that has started to go to school at 7 years or more. One needs to be

careful with this observation. Selection issues for parents able and willing to afford preschool are quite severe.

To summarize the results of the first stage, we can draw the following conclusions. Girls do relatively better in dictation and language and relatively worse in mathematics and general knowledge, when compared to boys. The performance of immigrant children is, in general, poorer than of Spanish students, controlling for the family background of the children. However, there are exceptions: Chinese students perform very well in mathematics; they get on average about one half of a standard deviation more than Spanish students and Romanian students perform, on average, better than Spanish students in all the subjects (except in general knowledge). We also find that starting the school at an early age is associated with better student's performance, although a causal interpretation is not necessarily warranted. The education and the profession of the parents appear among the most important determinants of the academic performance of children. If we compare the magnitude of the coefficients, the effect of the education of the parents is stronger than the effect of the profession.

#### **4.2. Second stage results**

To perform the second stage regression, we extract the coefficients  $\delta_j$  of the school dummies in the first stage regressions. They are estimates of the residual value of school on the performance of the children, once we control for individual and family characteristics of the students. We use these coefficients as the dependent variable in the second stage of our estimation procedure. They are regressed against all the variables at school level. Variables at school level are usually policy relevant variables. They reveal the correlations of different school characteristics with student performance.

The second stage estimations are weighted by the inverse of the estimated variance of the school fixed effects from the first stage.

Tables 4, 5 and 6 summarize the results of the estimations for the 2008/09 cohort for all the five subjects of the CDI exam. Because of the endogeneity of most of the control variables, we gradually add variables in our estimations in order to observe the sensitivity of the coefficients.

**Table 4**  
**School fixed effects regressed on school level variables (2nd stage - weighted regressions)**  
**2008/09 cohort**

	Dictation				Mathematics			
	OLS - All schools	OLS - All schools	OLS - Public schools	OLS - Public schools	OLS - All schools	OLS - All schools	OLS - Public schools	OLS - Public schools
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Class size	-0.004 (0.003)	-0.002 (0.003)	-0.012** (0.005)	-0.014** (0.006)	-0.002 (0.003)	-0.000 (0.003)	-0.002 (0.005)	-0.001 (0.005)
Enrollment 6th grade	0.002*** (0.001)	0.000 (0.001)	0.001 (0.001)	0.002 (0.002)	0.001** (0.000)	0.000 (0.000)	-0.001 (0.001)	-0.004** (0.002)
Private school	0.198*** (0.048)	0.015 (0.055)			0.191*** (0.043)	0.040 (0.050)		
Charter school	0.097*** (0.028)	-0.008 (0.032)			0.068*** (0.025)	-0.008 (0.028)		
School east	-0.019 (0.037)	-0.035 (0.038)	0.008 (0.047)	-0.069 (0.058)	-0.019 (0.032)	-0.015 (0.034)	0.047 (0.042)	0.081 (0.051)
School north	0.247*** (0.047)	0.199*** (0.047)	0.250*** (0.056)	0.193*** (0.063)	0.027 (0.041)	-0.009 (0.041)	0.082* (0.049)	0.106* (0.055)
School west	0.140*** (0.041)	0.091** (0.042)	0.177*** (0.055)	0.121* (0.068)	-0.051 (0.036)	0.099*** (0.037)	-0.003 (0.049)	0.023 (0.059)
School south	-0.006 (0.029)	-0.017 (0.031)	0.069* (0.041)	0.003 (0.051)	-0.012 (0.026)	-0.004 (0.028)	0.078** (0.037)	0.079* (0.045)
% Immigrant students in 6th grade (0-10%)		0.131** (0.051)	0.099* (0.059)	0.062 (0.071)		0.044 (0.045)	0.053 (0.052)	-0.015 (0.063)
% Immigrant students in 6th grade (11%-20%)		0.057 (0.048)	0.035 (0.054)	0.011 (0.064)		0.006 (0.042)	0.041 (0.048)	-0.037 (0.056)
% Immigrant students in 6th grade (21%-30%)		0.126** (0.049)	0.100* (0.055)	0.115* (0.064)		0.060 (0.043)	0.076 (0.049)	0.035 (0.056)
% Immigrant students in 6th grade (31%-40%)		0.072 (0.055)	0.040 (0.059)	-0.005 (0.068)		0.031 (0.048)	0.073 (0.053)	0.017 (0.059)
% Parent's education - Univ.		0.004*** (0.001)	0.004*** (0.001)	0.003* (0.002)		0.002*** (0.001)	0.002 (0.001)	0.001 (0.001)
% Parents' education - Higher secondary		0.001 (0.002)	0.001 (0.002)	-0.002 (0.002)		-0.001 (0.001)	-0.001 (0.002)	-0.004* (0.002)
% Parents' education - Vocational training		0.004** (0.002)	0.004* (0.003)	0.001 (0.003)		0.003* (0.002)	0.003 (0.002)	0.000 (0.003)
% Parents' education -Lower secondary		0.001 (0.001)	0.001 (0.002)	0.002 (0.002)		-0.002 (0.001)	-0.002 (0.001)	-0.001 (0.002)
% Parents' education - didn't finish compulsory studies		-0.004* (0.003)	-0.003 (0.003)	-0.003 (0.003)		-0.005** (0.002)	-0.006** (0.003)	-0.006** (0.003)
Teacher/student ratio			-0.143 (0.166)	-0.007 (0.216)			0.203 (0.149)	0.159 (0.191)
Av. age teacher			0.013*** (0.003)	0.012*** (0.004)			0.020*** (0.003)	0.023*** (0.004)
No. AMPA activ./total no. of activ.				0.242** (0.115)				0.176* (0.102)
% Students free meal				-0.002				0.001

				(0.003)				(0.002)
% Students with special educational needs				-0.001				-0.002
				(0.002)				(0.002)
School size				0.000				-0.000
				(0.000)				(0.000)
School size squared				-0.000				0.000
				(0.000)				(0.000)
Constant	0.058	-0.151	-0.491**	-0.375	-0.040	-0.081	-0.956***	0.878***
	(0.075)	(0.110)	(0.235)	(0.306)	(0.067)	(0.098)	(0.211)	(0.270)
Observations	1219	1219	728	558	1219	1219	728	558
R-squared	0.07	0.11	0.13	0.13	0.03	0.07	0.11	0.12

Notes:

1. Base categories for dummy variables: School capital, %Immigrant students in 6th grade more than 40%, Public school.
2. Robust standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

**Table 5**  
**School fixed effects regressed on school level variables (2nd stage - weighted regressions)**  
**2008/09 cohort**

	General Knowledge				Reading			
	OLS - All schools	OLS - All schools	OLS - Public schools	OLS - Public schools	OLS - All schools	OLS - All schools	OLS - Public schools	OLS - Public schools
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Class size	-0.000 (0.004)	0.002 (0.004)	-0.001 (0.006)	-0.001 (0.007)	0.002 (0.003)	0.004 (0.003)	-0.001 (0.004)	-0.002 (0.005)
Enrollment 6th grade	0.001 (0.001)	0.001 (0.001)	0.002 (0.001)	0.001 (0.002)	0.000 (0.000)	-0.001 (0.000)	-0.001 (0.001)	-0.003 (0.002)
Private school	0.056 (0.052)	0.020 (0.061)			0.145*** (0.041)	0.006 (0.048)		
Charter school	-0.077** (0.031)	-0.091*** (0.035)			0.047* (0.024)	-0.018 (0.027)		
School east	0.068* (0.040)	0.080* (0.042)	0.124** (0.055)	0.140** (0.067)	-0.040 (0.032)	-0.032 (0.034)	0.039 (0.042)	0.001 (0.052)
School north	0.094* (0.050)	0.093* (0.051)	0.109* (0.064)	0.132* (0.072)	0.129*** (0.040)	0.097** (0.041)	0.163*** (0.049)	0.146*** (0.055)
School west	0.039 (0.045)	0.025 (0.046)	0.109* (0.064)	0.093 (0.078)	0.106*** (0.036)	0.057 (0.036)	0.127*** (0.049)	0.086 (0.060)
School south	0.068** (0.032)	0.084** (0.035)	0.149*** (0.048)	0.139** (0.059)	-0.024 (0.026)	-0.006 (0.027)	0.083** (0.036)	0.053 (0.045)
% Immigrant students in 6th grade (0-10%)		-0.058 (0.057)	-0.064 (0.068)	-0.131 (0.082)		-0.046 (0.045)	-0.091* (0.052)	-0.115* (0.063)
% Immigrant students in 6th grade (11%-20%)		-0.053 (0.054)	-0.024 (0.063)	-0.064 (0.075)		-0.038 (0.042)	-0.064 (0.048)	-0.122** (0.057)
% Immigrant students in 6th grade (21%-30%)		0.029 (0.055)	0.029 (0.064)	0.044 (0.075)		-0.020 (0.043)	-0.054 (0.048)	-0.068 (0.057)
% Immigrant students in 6th grade (31%-40%)		0.050 (0.062)	0.029 (0.069)	0.017 (0.079)		-0.019 (0.048)	-0.017 (0.052)	-0.057 (0.060)
% Parent's education - Univ.		0.002** (0.001)	0.001 (0.002)	-0.000 (0.002)		0.003*** (0.001)	0.004*** (0.001)	0.001 (0.001)
% Parents' education - Higher secondary		0.000 (0.002)	-0.000 (0.002)	-0.001 (0.003)		-0.002 (0.001)	-0.002 (0.002)	-0.004** (0.002)
% Parents' education - Vocational training		0.001 (0.002)	0.004 (0.003)	0.001 (0.003)		0.002 (0.002)	0.003 (0.002)	0.002 (0.003)
% Parents' education -Lower secondary		0.001 (0.001)	0.001 (0.002)	0.003 (0.002)		0.000 (0.001)	-0.001 (0.001)	-0.000 (0.002)
% Parents' education - didn't finish compulsory studies		-0.003 (0.003)	-0.004 (0.003)	-0.004 (0.004)		-0.007*** (0.002)	-0.004 (0.003)	-0.004 (0.003)
Teacher/student ratio			0.311 (0.192)	0.584** (0.249)			0.037 (0.148)	-0.063 (0.192)
Av. age teacher			0.017*** (0.004)	* (0.005)			0.014*** (0.003)	0.014*** (0.004)
No. AMPA activ./total no. of activ.				0.100				0.112

				(0.133)				(0.102)
% Students free meal				0.003				-0.004*
				(0.003)				(0.002)
% Students with special educational needs				-0.002				-0.000
				(0.003)				(0.002)
School size				0.000				-0.000
				(0.001)				(0.000)
School size squared				-0.000				0.000
				(0.000)				(0.000)
Constant	-	-	-	1.520**				
	0.337***	-0.441***	1.266***	*	-0.223***	-0.238**	-0.760***	-0.397
	(0.081)	(0.122)	(0.274)	(0.353)	(0.066)	(0.097)	(0.208)	(0.271)
Observations	1219	1219	728	558	1219	1219	728	558
R-squared	0.02	0.03	0.05	0.08	0.04	0.07	0.11	0.09

Notes:

1. Base categories for dummy variables: School capital, %Immigrant students in 6th grade more than 40%, Public school.

2. Robust standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

**Table 6**  
**School fixed effects regressed on school level variables (2nd stage - weighted regressions)**  
**2008/09 cohort**

	Language			
	OLS - All schools	OLS - All schools	OLS - Public schools	OLS - Public schools
	(1)	(2)	(3)	(4)
Class size	-0.003 (0.003)	0.000 (0.003)	-0.008* (0.005)	-0.010* (0.006)
Enrollment 6th grade	0.001** (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.002)
Private school	0.184*** (0.045)	0.018 (0.052)		
Charter school	0.057** (0.026)	-0.032 (0.029)		
School east	-0.006 (0.034)	-0.012 (0.036)	0.048 (0.044)	-0.008 (0.054)
School north	0.222*** (0.043)	0.182*** (0.043)	0.239*** (0.052)	0.202*** (0.058)
School west	0.134*** (0.039)	0.085** (0.039)	0.184*** (0.052)	0.132** (0.063)
School south	0.009 (0.027)	0.010 (0.029)	0.109*** (0.039)	0.053 (0.047)
% Immigrant students in 6th grade (0-10%)		0.055 (0.047)	0.020 (0.055)	-0.029 (0.066)
% Immigrant students in 6th grade (11%-20%)		0.012 (0.044)	-0.002 (0.051)	-0.045 (0.059)
% Immigrant students in 6th grade (21%-30%)		0.081* (0.045)	0.054 (0.051)	0.064 (0.059)
% Immigrant students in 6th grade (31%-40%)		0.057 (0.051)	0.030 (0.056)	-0.014 (0.063)
% Parent's education - Univ.		0.004*** (0.001)	0.004*** (0.001)	0.002 (0.002)
% Parents' education - Higher secondary		0.000 (0.001)	-0.000 (0.002)	-0.003 (0.002)
% Parents' education - Vocational training		0.003* (0.002)	0.005* (0.002)	0.001 (0.003)
% Parents' education -Lower secondary		0.001 (0.001)	0.001 (0.002)	0.002 (0.002)
% Parents' education - didn't finish compulsory studies		-0.005** (0.002)	-0.004 (0.003)	-0.004 (0.003)
Teacher/student ratio			0.008 (0.156)	0.140 (0.201)
Av. age teacher			0.017*** (0.003)	0.018*** (0.004)
No. AMPA activ./total no. of activ.				0.216**



				(0.107)
% Students free meal				-0.002 (0.002)
% Students with special educational needs				-0.002 (0.002)
School size				0.000 (0.000)
School size squared				-0.000 (0.000)
Constant	-0.110 (0.070)	-0.278*** (0.102)	-0.880*** (0.222)	-0.773*** (0.284)
Observations	1219	1219	728	558
R-squared	0.06	0.10	0.14	0.13

Notes:

1. Base categories for dummy variables: School capital, %Immigrant students in 6th grade more than 40%, Public school.
2. Robust standard errors in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

In columns (1) and (2) of Tables 4, 5 and 6 we estimate our empirical model for all types of schools of the Spanish education system– public, private and charter. In these estimations we can only control for class size, enrollment in the 6th grade, the geographical location of the school, the share of immigrant children in the 6th grade and the shares of children whose parents have a certain level of education. In columns (3) and (4) we estimate the model only for public schools.

We find that private and charter schools perform significantly better than public schools and the difference between private and public schools is particularly large. Moving a 6th grade class from a public to a private school might increase its performance in all the subjects (except general knowledge) from the median to approximately the 70th percentile. However, the large positive and significant effect that we obtain for private and charter schools in the first column might be due to other school characteristics or parental characteristics that we do not account for in these estimations. In column (2), once we control for the socioeconomic composition of the classroom (parental education in the 6th grade) these differences disappear. A likely explanation for this finding is that this type of schools have been good at separating the children of better educated parents from the rest, and allowing their students to extract the “peer effects” of their classmates, which as we discuss later are sizable.

The second stage controls also for the share of immigrants in the class. Since there is a lot of variability in the number of immigrants among the schools and the effect of the number of immigrants could be nonlinear, we adopt the following strategy: we divide our sample of schools in five intervals, according to the percentage of immigrants in the total number of children in the 6th grade. The first interval includes schools with 0 to 10% of immigrants in the 6th grade, the second interval includes schools with 11% to 20% and so on, and the last interval includes schools with more than 40% immigrants in the 6th grade. In general, we do not find empirical evidence that the number of immigrants in a class is associated with the performance of children in that class, controlling for the average level of parental education in the 6th grade among other school level variables. We do find a statistically (weakly) significant effect of the number of immigrants in a class in the case of dictation. In particular, the strongest effect that we find is that a school that has between 21% and 30% immigrants in the 6th grade might improve the performance of the children in dictation from the median to the 65th percentile, with respect to a school that has more than 40% immigrants in the 6th grade. This could be interpreted as a negative effect of the share of immigrants in the class. However this result is not robust for the other subjects of the C.D.I. exam.

We also find that the average parental education in the class has a statistically significant impact on the school fixed effect. In particular, the estimations show that the percentage of children whose parents have university studies has a positive and significant effect on the performance of the children. That is, beyond the effect of their own parental background, the children experience an additional effect coming from their peers' parental background.

Two of the most important policy variables are the class size and the teacher/student ratio. They are in the centre of the public and academic debate since they are among the most important factors that increase schooling costs. Consistently with the observations from previous research we summarized in the literature review, we do not find any strong evidence that decreasing class size improves the academic performance of children. In dictation and language, the coefficient of class size is statistically different from zero, but only for public schools and, moreover, the magnitude of the effect is extremely small: reducing class size by 5 students would increase the average performance of children in the class by about 0.16 standard

deviations in the case of dictation and 0.13 in the case of language. This would improve the performance of the class in both dictation and language from the 45th percentile to the median. The coefficients for class size are not statistically different from zero in the rest of the subjects, once we properly control for other school level controls. The teacher-student ratio is not a relevant factor neither. Except in general knowledge, the coefficient of this variable is not statistically different from zero in any of the subjects.

The average age of the teachers in the 6th grade is statistically significant and it has a positive effect: the older the professor, the better the performance of the students in his class. However, the magnitude of the effect is quite small: increasing the age of the teachers by 5 years would improve the performance of a class from the 45th percentile to approximately the median in dictation and from the 65th percentile to the 75th percentile in both mathematics and language. A possible reason for this finding is that in Spain more senior teachers have priority when choosing schools. Thus, they may be able to pick better schools in dimensions which our observables miss.<sup>11</sup>

An interesting result that we derive from the estimations for the 2008/09 cohort is that extracurricular activities matter. More concretely, we find that the ratio of the total number of extracurricular activities organized by Parents' Associations to the total number of extracurricular activities in the school has a positive and statistically significant effect on the results in dictation, mathematics and language. The performance of a class in any of these three subjects in a school where half of the extracurricular activities are organized by Parents' Associations would move from the median to a bit more than the 60th percentile. A reasonable interpretation is that this variable reflects groups of especially active and dedicated parents which our observables cannot capture.<sup>12</sup>

To summarize, the second stage results show little empirical evidence that differences in resources among schools affect student outcomes, once individual and

---

<sup>11</sup> In a separate specification, we used the age of the youngest teacher in school instead of average age, since younger teachers are probably not able to choose school. It had a positive sign, but it was not significant in most of specifications.

<sup>12</sup> We should mention that since not every school has a Parents' Association, it is possible that the effect that we see comes from the fact of having one and not from the activities itself. Our data is on extracurricular activities and who organizes them and we do not have information on whether a school has a Parents' Association or not. We could infer that schools which have a 0% of extracurricular activities organized by Parents' Associations do not have one. We included this dummy variable in our estimations, however it was only significant in Mathematics and only at 10% level. This might be due to the fact that we do not have precise information on which schools have Parents' Association.

family background characteristics are accounted for. In particular, we find no empirical evidence that reducing class size or increasing teacher/student ratio would generate improvements in the academic performance of the students from primary schools. However, we do find that extracurricular activities organized by Parents' Associations are associated with better student achievement, which is consistent with a positive effect of the involvement of parents in the education of their children. This result strengthens the effect that we previously found for the average parental education in the class.

### **4.3 Robustness checks and additional results**

In this section we perform a series of robustness checks and we provide some additional results that might shed light on our conclusions from the main empirical model.

We estimate the first stage allowing for interaction terms between gender and the education of the parents. Since all the results are robust to the inclusion of these interaction terms, we will only comment the differences that appear by gender and parents education. Table 7 includes the estimated coefficients of the interaction terms.

We still observe differences between boys and girls, when we compare boys and girls with the same level of education of the parents. Comparing boys and girls with parents with a high level of education (University, Higher secondary and Vocational training), we find that girls get on average 0.2 standard deviations more than boys in dictation and about 0.1 standard deviations more than boys in language. Furthermore, boys whose parents are more educated perform better than girls whose parents have a similar level of education in mathematics and general knowledge: they get about 0.14 standard deviations more in mathematics and about 0.18 standard deviations more in general knowledge. When we compare boys and girls with parents with a low level of education (Lower secondary or didn't finish compulsory studies), we find the same differences. These are actually the differences that we detect in the main estimations, where we don't allow for the differences between boys and girls to depend on the level of education of parents.

Table 8 shows the results of the estimation of the first stage where we allow for the differences between girls and boys to depend on the nationality of the children. All

the results are robust to the introduction of the interaction terms between gender and nationality. Therefore, we only comment the estimated coefficients of the interaction terms.

Some interesting differences arise between boys and girls of the same nationality. Chinese boys do relatively better than Chinese girls in mathematics (about 0.3 standard deviations more), while they do relatively worse in dictation, reading and language (about 0.3 standard deviations less). Romanian girls perform relatively better than Romanian boys in dictation (0.3 standard deviations more) and language (0.2 standard deviations more), but they get about 0.1 standard deviations less in general knowledge. We already pointed that Latin American students is the group that performs relatively the worst and we find here the same result: all the coefficients for the dummy variables for Latin American boys and girls are negative and highly significant. When we compare Latin American boys and girls, we also observe statistically significant differences: Latin American girls perform relatively better than Latin American boys in dictation (about 0.2 standard deviations more) and in language (about 0.1 standard deviations more), but they perform relatively worse in mathematics and general knowledge (about 0.2 standard deviations less). Lastly, the performance of Moroccan girls and boys is similar in general in all the five subjects.

Finally, we estimate the specification from the first stage of our empirical strategy separately by type of school: public, charter and private. The main conclusions that we draw in section 4.2 do not change.<sup>13</sup>

---

<sup>13</sup> These results are available upon request.

## 5. Concluding Remarks

In this paper we have analyzed the data from a standardized exam in the region of Madrid. Previous research has documented an important effect of parental background on school outcomes. We also find that parental education and profession are essential predictors of students' results. But some of our results about parental background are intriguing. Spain has experienced a large inflow of immigrants from various nationalities in the past decade.<sup>14</sup> After controlling for their parental background, students of some nationalities do significantly better (and some do worse) than the local population. This needs to be analyzed further, but it could give insights into the relative importance of the organization of the educational system versus the social values those recent immigrants carry with them. Other data, like those coming from international exams (e.g. the PISA test) cannot separate the two effects so easily.

With respect to school level variables we find, like others before, that class size has at best a small effect on performance. But again, some results point to interesting further research. One of them is the effect of parental involvement in the school. We find that the percentage of school activities organized by the Parents' Association has a large and significant effect on the dictation, language and mathematics exams. It would be nice to understand whether this is simply because good parents are more active, or whether there is some interaction between parents, teachers and school administrator

---

<sup>14</sup> We went from less than 1 percent of immigrants in the population to almost 10 percent in that period.

## 6. References

- Angrist, J.D. and V. Lavy (1999), "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement", *The Quarterly Journal of Economics* 114, 533-575.
- Ermish, J. and M. Francesconi (2001), "Family Matters: Impacts of Family Background on Educational Attainments", *Economica* 68, 137-156
- Fertig, M. (2003), "Who's to Blame? The Determinants of German Students' Achievement in the PISA 2000 Study", IZA Discussion Paper No. 739
- Finn, J.D., and C.M. Achilles, "Answers and Questions about Class Size: A Statewide Experiment", *American Educational Research Journal* 27, 557-577.
- Hanushek, E.A., S.G. Rivkin, and L.L. Taylor (1996), "Aggregation and the Estimated Effects of School Resources", *The Review of Economics and Statistics* 78, 611-627
- Hanushek, E.A. (1999), "The Evidence on Class Size", in Susan E. Mayer and Paul Peterson (ed.), *Earning and Learning: How Schools Matter* (Washington, DC: Brookings Institution), 131-168.
- Hanushek, E.A. (1999b), "Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects", *Educational Evaluation and Policy Analysis* 21, 143-63.
- Hanushek, E.A. (2003), "The Failure of Input-based Schooling Policies", *Economic Journal* 113, F64-F98.
- Hoxby, C. M. (2000), "The effects of class size on student achievement: New evidence from population variation", *Quarterly Journal of Economics* 115, 1239-1285.
- Krueger, A.B. (1999), "Experimental Estimates of Education Production Functions", *Quarterly Journal of Economics*, 114, 497-532.
- Leuven, E., H. Oosterbeek and M. Rønning (2008), "Quasi-experimental Estimates of the Effect of Class Size on Achievement in Norway", *Scandinavian Journal of Economics* 110, 663-693.

Loeb, Susanna and John Bound (1996), “The Effect of Measured School Inputs on Academic Achievement: Evidence from the 1920s, 1930s and 1940s Birth Cohorts”, *The Review of Economics and Statistics* 78, 653-664

Woessman, L. (2004), “How Equal Are Educational Opportunities? Family background and student achievement in Europe and the US”, CESIFO Working Paper No. 1162