Does money account for regional differences in student performance in Spain?

Yalina Zaldivar[1]

Supervisor: Katie Baird[2]

Departamento de Fundamentos del Análisis Económico II

Universidad del País Vasco

12 /09/2008

[1] E-mail: yalinazaldivar@hotmail.com

[2] Associate Professor of Economics. Interdisciplinary Arts and Sciences University of Washington, Tacoma. E-mail: kebaird@u.washington.edu

INDEX

1. INTRODUCTION

Education is an area with crucial and long-lasting consequences, which affects the economy and the future social capital of a country. Education should also be a way to ensure equity of opportunities for the society. However, in the case of Spain, there can be some possible doubts about the last mentioned function of education, since there seem to be disquieting differences in education among its regions. These doubts were reinforced with the PISA 2006 results, in which 10 of the 17 autonomous communities were examined and in which a lack of homogeneity in students' results was revealed (Table 1).

Castile and Leon and La Rioja are the best positioned communities in Spain, both share the 4th place in science for the PISA ranking (when including all the 57 participating countries), actually in a better position than Germany or the United Kingdom (Lacasa,2008). When comparing these communities with other Spanish communities, their better performance in PISA 2006 is remarkable. On the contrary, Andalusia, Catalonia and the Rest of Spain[3] obtained the lowest results in Spain although they still remain above the science average of the 57 participating countries. Several factors might be responsible for these considerable differences within Spain.

Even though it can be supposed that money plays an important role in education quality, we may not presume that it is the factor that conditions the students' results. In Spain, the per pupil expenditure (PPE) differs largely from community to community. There are six communities with a higher PPE than the country mean of 4911 Euros per student, the Basque Country is at the top (7708) and Cantabria (4096) at the bottom of the list. Trying to explain high test scores as a result of high PPE may not be suitable especially for communities as the Basque Country, Navarre and Asturias who have the three highest PPE but got the 8th, 3rd and 5th place in the

[3] The Rest of Spain refers to: Canarias, Castile-La Mancha, Extremadura, Illes Balears and Murcia. Madrid and Valencia did not present the examination.

science score. On the contrary, Castile and Leon and La Rioja are the communities with the 4th and 5th highest PPE but got the first two places on PISA science scores (see Table 1 and 2). Following this thought, it does not show up why La Rioja and Catalonia got such good results if their PPE investment per student is not as big as others.

The aim for equality of educational opportunities and policy makers who would like to know how to achieve a high productivity of schooling have inspired decades of research on school effects by the impact of school resources. However, the availability of resources that can be bought with money is considered with different impacts in the literature.

A first group of authors conclude that resources are related to student achievement in school. Better equipment, better prepared and educated teachers, reduced class-sizes as well as computer availability at school were associated in the literature with higher achievement. However, no consensus was reached about the precise role of these money related inputs on the students output.

Moreover, some argue that Money Related Variables (MRV) are not a decisive factor for students results. The non-money related variables that are associated with student achievement in school can be subdivided in two categories, the student characteristics on one hand and the school characteristics that are not money related on the other hand.

Student characteristics that are shown in several studies to be related with educational performance are: the educational level of the parents, parental occupation, the number of books at home, gender, repetition and delay in studies and migration background.

Some studies in the literature deduce that differences in schooling institutions matter much more for international differences in student performance than differences in educational expenditures and conclude that educational policy in Europe should not focus on providing more resources to schools, but should focus rather on improving the institutional environment in which schools function.

The contribution of this thesis is to estimate the extent to which the average performance difference of science scores can be due to differences in the average value of Money Related Variables (variables whose values may change according to the amount of money addressed to education) between regions in Spain, oppositely to Non MRV (factors that are unrelated to school spending). Using a function to measure the relation between the explanatory variables (money and not money related variables from PISA 2006) and the science test scores obtained by each region, this paper examines the extent to which regional differences can be explained by the differences in the average value of Money Related Variables.

In the next chapter, a literature review will give an overview of the studies made in this field, highlighting the relevance of MRV and NMRV. Chapter 3 gives an overview of the PISA data that are used in this thesis and gives an explanation of the data analysis technique and the empirical model used to obtain the desired results. The general results obtained by using the model are discussed in the first part of chapter 4 while the second part is emphasized to the connection between science test scores and regional spending. A last chapter includes a discussion and a summary of this study.

2. LITERATURE REVIEW

Since education is key to individuals' and societies' future social and economic performance, policy makers would like to know how best to achieve a high educational performance. A number of studies have tried to identify the (economic) factors that might be responsible for differences in student performance.

Gamoran and Long stated in their compilation "Equality of Educational Opportunity: A 40-year retrospective" (2006), that equality of educational opportunities has inspired decades of research on school effects by the impact of school resources. Arguing about the money related inputs in school and their effects on the students output, availability of resources that can be bought with money is considered with different impacts in the literature. Many studies have been developed with education production functions in order to attempt to solve the students' input – output puzzle.

A first group of authors contribute to the puzzle in the literature on educational production functions by trying to demonstrate a positive impact of increased spending on educational outcomes.

Fuchs and Woessmann (2004b) used the PISA student-level achievement database to estimate international education production functions. They concluded that resources are significantly related to science achievement. In particular, they found that while smaller classes do not go hand in hand with superior student performance, better equipment with instructional material and better educated teachers do. Another conclusion obtained was that resources contribute considerably to the international variation in student performance.

Mora, Escardibul and Espasa (2005) studied the effects of educational policy on educational performance at a regional level in Spain. Specifically, they considered the effects of three policy instruments (expenditure per pupil, class-size and pupil teacher ratio[4]) on three educational outcomes related to academic failure in schools: regional dropout rates at the end of compulsory education (at age 16), the percentage of students who were required to repeat one or more academic years (at age 15), and the percentage of students who fail the university entrance examination. Their results and thus the implications for policymakers are simple: educational policies addressed to give further attention on pupils (reducing class-size or the pupil-teacher ratio) have higher effects on dropout and repetition rates than expenditure per pupil. They concluded that policies aiming to improve a student's personal attention seem to succeed in reducing school failure rates before the end of compulsory education, while higher expenditure per pupil helps to increase a student's performance after compulsory education.

Fuchs and Woessmann (2004a) estimated the relationship between computers and students' educational achievement in the international student-level PISA database. Bivariate analysis showed a positive correlation between achievement and computer availability at school. Achievement showed an inverted U-shaped relationship with computer and internet use at school.

Greenwald, Hedges and Laine (1996) carried out reviews of the literature finding an effect of school resources whereas Hanushek (1997) did not find a persistent effect of school resources. The differences between these studies are based primarily on their different criteria for including studies in their meta-analyses, which resulted in different summaries of results. Greenwald et al. were much more selective, excluding or down weighting studies when multiple findings derived from the same data. They also took into account the magnitude and variability of effects across studies, whereas Hanushek tallied positive, negative, and neutral findings. As a result of their

[4] Class size was defined as the average number of students per educational unit in E.S.O. whereas pupil – teacher ratio considered the average number of students per teacher in primary and secondary education (there was no data available only for secondary education). Though both policies did not address the same inputs (given that class-size considers classroom inputs other than just the number of teachers), they were included as alternative factors in the estimations.

more selective scrutiny and synthetic approach, Greenwald et al. found moderate effects of school resources such as teacher's salaries and smaller class size. Despite their differences, Greenwald et al. and Hanushek agreed on three points: (a) in at least some cases, higher level resources are associated with higher achievement; (b) the qualities of schools that produce these effects are hard to pin down; and (c) the way in which resources are used is more consequential for achievement than the presence or absence of resources.

On the other hand some argue that Money Related Variables (MRV) are not a decisive factor for students results.

The Coleman report in 1966 discovered that differences in average resources among schools were not nearly as great as expected and that the impact of school resources on student achievement was modest compared to the impact of student's family background.

Woessmann (2005) views the process of knowledge creation in schools as a production process, where several 'inputs' are combined to create an 'output', namely the cognitive skills of students. The inputs he used included factors as students of a given background, teaching materials, teachers of a given educational background and class sizes. This paper estimated the education production function using representative samples of middle-school students in 15 West European countries. The results of this paper suggest that expensive reduction of class sizes is extremely unlikely to foment student learning, what was also concluded in an anterior paper of Fuchs and Woessmann (2004b).

In his paper, Gamoran et al. (2006) mentioned Montagnes (2001) who confirmed that studies of school effects conducted in developed countries (like Spain) probably say less about the effects of school resources than in developing countries, since the poorest schools in developed countries still have many more resources than the typical school in developing ones-.

Although there is a lot of discussion about the role of money related inputs in school on the effects of the students output, more consensus is reached about the role of variables which are not related with money.

A first group of the non-money related variables include the student characteristics and the family backgrounds of the students.

Fuchs and Woessmann (2004b) concluded in their paper already mentioned above that student characteristics, family backgrounds and home inputs are significantly related to science achievement. In particular, they found that students' family background -measured by parental education, parental occupation or the number of books at home- is consistently strongly correlated to their educational performance and that boys outperform girls in science.

On its PISA 2006 report, FETE-UGT (2007) concluded that the relationship between the educational level of the parents and the results of the students is evident. Their results show that students whose parents did not finalize the compulsory studies score less than students whose parents finalized university studies. The educational level of the mother followed by the number of books at home are, according this study, two of the factors which affect the most the results of the students. Another conclusion that was made is that differences in science achievement are not always favored for a specific gender. As was concluded in previous studies, they identified that the repetition and the delay in studies is not beneficial for the academic results for the students in the acquisition of the basic competences. Their results also indicated that the scores of students that are born outside Spain are on average lower than those born in Spain.

The High-Priority Indicator Report from MEC (Martín, 2007) shows that E.S.O. students with at least one parent with university studies (ISCED 5) perform better than those whose parents did not achieve university level (ISCED 4 and lower). This report also shows that male students and students from private schools yield better results than their counterparts in Science, Geography and History – see Figure 1-.

Some argue that families from continents as Africa and America[5] show a lack of attention and expectations for their children's education and do not give them the same importance and support for their studies as native families (Goenechea, 2006). According to the OMBUDSMAN's School Violence Report among equals in E.S.O. (UNICEF, 2007), one of the most relevant factors that may affect the performance of the immigrant students is their high propensity of being victims of school violence (which mostly takes place in the last years of primary and in secondary courses), mainly by being ignored or threatened by colleagues for cultural or religious matters. Figure 2, shows that Balears (part of the "Rest of Spain") and Catalonia are communities with a high foreign population and poor science scores. Surprisingly, Navarra and La Rioja are also among the communities with the highest student foreign population but are also located among the first places of PISA science scores (4th and 2nd place respectively). Furthermore, Andalusia does not show a high foreign student population at E.S.O.

This report also found that violence – ignore, destructive critics and gossips - among students from private schools is significantly higher than those of public schools. Regarding the gender, the time spent by males affecting colleagues from both genders by not allowing them to participate, establishing nicknames, hiding or breaking other student's goods, threatening and sexually harassing is significantly higher. Females are significant authors of curses but spend just a little part of their school time on it.

Also Entorf and Lauk (2007) analyzed the school performance of migrants dependent on peer groups in different international schooling environments by using data from the international OECD PISA test. Their results suggest that non-comprehensive school systems seem to magnify the already existing educational inequality between students with a low parental socioeconomic migration background and children from more privileged families and that student with a migration background and a disadvantageous parental status would benefit from higher diversity within schools.

---

[5] South America and Africa are the origin countries of the majority of foreign student population in Spain. (MEC, 2007).

A second group of the non-money related variables include the school characteristics that are not money related.

Gundlach and Woessmann (2001) concluded that an obvious problem for education policy is that spending more money on schools did not have any payoff in Europe. Their results showed that differences in schooling institutions matter much more for international differences in student performance than differences in educational expenditures. From these results it can be suggested that educational policy in Europe should not focus on providing more resources to schools, but should focus rather on improving the institutional environment in which schools function, since spending more money within an institutional system that sets poor incentives will not improve student performance.

Fuchs and Woesmann (2004b) concluded that institutions of the school system are important for student achievement. External and standardized examinations seem to be performance-conducive. The effect of school autonomy depends on the specific decision making area. School autonomy is mostly beneficial in areas with informational advantages at the local level (process and personnel decisions), whereas it shows detrimental effects in areas that are prone to local rent-seeking activities (setting of standards and budgets). Private school operation seems to be conducive to student achievement, while at the same time public school financing.

3. DISCUSSION OF DATA AND DATA ANALYSIS TECHNIQUE

3.1. General discussion of PISA data

The dataset used in this analysis is the Programme for International Student Assessment (PISA). PISA is an internationally standardised assessment that was jointly developed by participating countries and administered to 15-year-olds in schools. The survey was implemented in 43 countries in the 1st assessment in 2000, in 41 countries in the 2nd assessment in 2003 and in 57 countries in the 3rd assessment in 2006. In all PISA cycles the domains of reading, mathematical and science literacy are assessed. These subjects are evaluated in each assessment but every three years one of the subjects is rotated and receive deeper attention - in the sense that it includes an extensive set of questions related to this domain-. For the PISA 2006 cycle, the focus was on scientific literacy.

In addition to assessing student performance, PISA also collects information on student attitudes and approaches to learning as well as the learning environment and organisation of schooling. These files include questionnaires, codebooks and data files in ASCII format in order to process the data. The student questionnaires of 2006 give information about the specific student, his family, his views on various issues related on science, the environment, careers, learning time, teaching and learning science. The information of the school questionnaires of Spain of 2006 helps to illustrate the similarities and differences between groups of schools in order to establish the impact of resource distribution on student achievement within Spain. For the PISA 2006 assessment, there is information available for autonomous communities that asked for an extended sample such as: Andalusia, Aragon, Asturias, Cantabria, Castile and Leon, Catalonia, Galicia, La Rioja, Navarre and the Basque Country. In this thesis, available data are combined to construct a dataset containing 23266 students in Spain who took the test.

3.2. Data

The dataset combines average student test scores in science with student background characteristics, teacher characteristics, classroom characteristics and school characteristics. In these groups of variables, a subgroup of variables is formed according to the ones "money can buy", the Money Related Variables (MRV). For estimation purposes, a variety of the qualitative variables were transformed into dummy variables. For each variable, there will be some missing values. Therefore, the used dataset is not complete. However, since the amount of original versus missing data for each variable is low, neglecting them will not cause significant changes in the quantitative analysis. Table 3 gives an overview of the variables used in this thesis and presents their descriptive statistics (PISA data from school and student questionnaires). All these variables are described below.

3.2.1. Dependent variable: science test scores

In this thesis only the science test score is analyzed. The reason for this is twofold. First, because the main focus of the PISA 2006 study was on science. Second, since the correlation between the three subjects is substantial (0,7911 between science and reading, 0,8750 between science and math and 0,7541 between reading and math) looking at one test can be justified. An average science test score is calculated from the 5 plausible test scores of science.

3.2.2. Independent variables

The independent variables can be categorized in money-related-variables (MRV) and non money-related-variables (NMRV). On the one hand, MRV are the ones which "money can buy", i.e. variables that could change if only the school had more money. These variables can describe teacher characteristics, classroom characteristics or other school characteristics. On the other hand, there are the NMRV which can be subcategorized in Student Background Variables and Non-Money-Related School Variables.

3.2.2.1. Money-related-variables (MRV)

The first variables considered in this section describe the share of teachers in the schools that have a proper certification or qualification and the lack of qualified science teachers. As there is no way with PISA data to relate a student with its own teacher, this variables give an approximation of the characteristics of the teacher a student may have on average in his school. Consequently these variables are imperfect measures. It is true that even though a school has a lack of qualified science teachers, a specific student can have a qualified science teacher. However the probability that this student has a qualified science teacher will be lower.

It may be predicted that teachers, fully certified by the appropriate authority, do in general better than teachers who do not have this certificate, leading to better general results of schools with a higher percentage of certified teachers. This variable can be considered as a Money Related Variable since certified teachers may gain more than teachers without an appropriate certification. Therefore, schools with a higher percentage of certified teachers possibly have to spend more than schools with a low percentage of certified teachers.

A similar explanation can be given for teachers with an ISCED5A qualification. Also here it is expected that students of schools with teachers with on average a higher level of education (a higher percentage of teachers with an ISCED5A qualification) will perform better than otherwise. This variable can be considered as a Money Related Variable since the opportunity cost of teaching (i.e. science) will be too high for certified teachers relative to other professions[6].

[6] Teachers from Andalusia, Galicia, Ceuta and Melilla (all regions with poor test score results) are the ones with the lowest wage in the Spain. The lack of wage homologation between the communities makes possible a difference in Andalusia of € 230 per month below the mean of what a teacher earns in other places such as: Castile-La Mancha, Cantabria, Navarra and the Basque Country. Just Galicia, Ceuta and Melilla offer a lower wage to their teachers than in Andalusia. (Junta de Andalusia, 2007).

Another indicator is the lack of qualified teachers, in this case the science teachers. Using the questionnaire answers by school principals, two dummies are created for this indicator. On the one hand, a dummy is created for schools where there seems to be not at all or just a very little lack of science teachers. On the other hand, a dummy is created for schools with some extent or a lot of lack of science teachers. For this variable, the dataset is not complete since around 5% are missing values, that will be included in the reference group later on. A lack of science teachers means that some students do not receive classes from a qualified science teacher and are thus supposed to do worse than otherwise. Also this variable can be seen as money related, since more money can buy more qualified science teachers. If science teachers were well paid, that would certainly motivate the most skilled persons with an appropriate degree to become teachers.

The main money related indicator describing the classroom characteristics is the average class size. A small class size is assumed to have a positive effect on student performance, since more attention can be given to each individual student. Four dummies are constructed to quantify the average class size, going from less than 15 students, from 16 up to 30, from 31 until 50 and the last dummy for classes with 51 or more students. Around 5% of the dataset for this variable are missing values. The more money spent, the smaller the classes can be, because smaller classes means more costs to pay (i.e. more teachers, more classrooms and thus space used, more material for each classroom, etc.). Smaller classes may mean that more attention goes to each student from the teacher which may positively affect output.

A school related variable that is expected to be positively related to the student performance is the computer-student ratio. The more useful resources a student receives the better will be the test scores. This variable can be seen as money related since more resources available in a school mean that more money was spent by the school on technology.

In terms of material inputs, questionnaire answers by schools can be used on how much learning in their school is hindered by lack of instructional material, such as science laboratory equipment, instructional materials, computers for instruction, Internet connectivity, computer software for instruction, library materials and audiovisual resources. Two dummies are created: on the one hand, a dummy is created for schools that have not at all or very little lack of all instructional material. On the other hand, a dummy is created for schools that have to some extent or a lot of lack of all instructional material. The reference group mostly contains schools laying in between these two extremes and a small group of missing values. Students in schools whose principals reported not being hindered by a lack of material are supposed to perform better than students in schools whose principals reported insufficient instructional material.

3.2.2.2. Non-money-related-variables (NMRV)

- Student Background - NMRV

The grade parameter is subdivided in 5 dummies, going from 7th grade to the 11th grade. It can be expected that educational performance of 15-year-old students increases with the grade level in which they are enrolled. After all, since PISA is administered to 15-year-old students, this parameter captures if students have repeated or skipped a grade. A 15-year-old student attending classes in the 7th grade has already repeated a grade, while a 15-year-old student attending classes in the 11th grade has already skipped a grade.

Due to the difference in gender composition in the communities, dummies named female and male were created to include a possible gender effect on test scores. Since in this work the focus will just be given to science score, a possible outcome can be that females perform worse in science than males.

Immigration status of students and their families can be related to their educational achievement. It is expected that native students perform better than first generation immigrants who on their turn are expected to do better than second generation immigrants. Therefore, these 3 categories are included as dummies, while around 1% of the dataset for this variable are missing values.

Another category of variables that can have an effect on the achievement on the test concerns the educational level of the parents. Instead of using both the educational level of the father and the mother, taking only the highest educational level of the parents will be meaningful and may influence the achievement in school of their child since a parent with a high level of education may create an environment that stimulates the child to perform well in school, i.e. the parental support will be higher. Three Dummies are created in this category: both low educated (ISCED level 0, 1 or 2), at least one parent with a medium level of education (ISCED level 3 or 4), or a high level (ISCED level 5 or 6). The ISCED level stands for: ISCED 0 (Pre-primary education), ISCED 1 (Primary education), ISCED 2 (Lower-secondary education), ISCED 3 (Upper-secondary education, also including TVET -Technological and Vocational Educational Training-), ISCED 4 (Post-secondary non-tertiary education), ISCED 5A (B.A, B.Sc., M.A, M.Sc., M.B.A, M.D, D.M.D), and ISCED 5B (B.Ed. etc.), and ISCED 6 (PhD, Dr.Sc.). Around 2% of the dataset for this variable are missing values.

A category of variables that can be related with the category mentioned above is the highest parent White/Blue collar classification, since a distinctive element of blue-collar work is the fewer requirements for formal academic education which is needed. The colour-coding can been used to identify a difference in socio-economic class and the time available to support school assignments. Both groups can be subdivided in high and low skilled, what gives 4 dummies in total. Around 2% of the dataset for this variable are missing values. It is expected that students with a white collar and high skilled parent will score best, while students with both blue collar and low skilled parents will score worst. Also here, the difference in parental support can be a possible explanation for that.

Another indicator of the student family background that can be related to his performance is the number of books at home. It can be expected that the effect is larger for reading scores; nevertheless, there may also exist an effect on science. Three dummies are created, one where students report from zero to 25 books in the home, from 26 to 100 (moderate number of books) and more than 101 books, considered as a large number of books available at home. Around 2% of the dataset for this variable are missing values.

As a last measure of student background characteristics, the time spent per week in science can be considered. This time adds the time spent attending regular lessons at school, attending out-of-school-time lessons and studying or doing homework by him/herself. Since the focus of this study will be on the performance of science, just the time in this subject is considered. Two dummies were created, one for students who spend less than two hours in both regular lessons, out-of-school-time lessons and homework, and one for the ones who spend more than two hours per week in each category, both the extreme cases of devoting a little or a lot of time in science. The reference group mostly contains students laying in between these two extremes and a small group of missing values and covers around 11% of the dataset for this variable. The bigger the time spent on science may explain better students' results.

- School Characteristics - NMRV

A first school related categorisation is the public versus private schools. For each group, a dummy is made. A lot of research is already done to search for a difference in quality between two groups. Around 5% of the dataset for this variable are missing values. However, there is no agreement if this variable matters or not.

Another possible cause for performance differences may be the location of the school. There may exist a performance advantage of schools in cities relative to village schools, since schools in cities may be more accessible and thus preferred by teachers to go to. To include this possible effect, two dummies were created for different categories of community locations. On the one hand a dummy is created for schools located in a village hamlet, a rural area, or a town while cities are grouped in the second dummy. Around 4% of the dataset for this variable are missing values.

3.3. Data analysis technique: The Empirical Model

As mentioned above, the average science test score is chosen as dependent variable, while the independent variables can be categorized in Money-Related-Variables (MRV) and Non Money-Related-Variables (NMRV). The NMRV are subcategorized in Student Background Variables and Non-Money-Related School Variables.

These groups of variables can be used in a country wide education production function for Spain. The education production function in this case is a function that maps quantities of measured inputs to school (money related characteristics and characteristics not related with money) to a measure of school output, the science test score. Analogous to similar studies in the past and for empirical purposes one might assume that this function is linear, i.e. the relationship between the independent variable and the dependent variable is modeled by a least squares function. This function is a linear combination of the regression coefficients. The results are subject to statistical analysis. The micro econometric estimation equation of the country wide education production function will have the following form:

$$Y = MRV \cdot \beta_1 + NMRV \cdot \beta_2 + \varepsilon$$

Where Y stands for the science test scores, MRV is a vector of the money related variables, $NMRV$ is a vector of the non-money-related variables and ε is the disturbance term; the parameter vectors β_1 and β_2 will be estimated in the regression. The vector of non-money-related variables can be subcategorized in student background and school $NMRV$, giving:

$$Y = MRV \cdot \beta_1 + St \cdot \beta_2 + Sc \cdot \beta_3 + \varepsilon$$

Where St stands for the vector of student background variables and Sc the vector of non-money related school variables.

This specification of the Spanish education production function restricts each effect to be the same in all regions, as well as at all levels (within schools and between schools) and thus excludes the possibility of analysing potential heterogeneity of certain effects between regions and between levels.

To obtain representative estimates, weighted least squares (WLS) estimation is used, using the given student weights of PISA as probability weights. These probability weights, or sampling weights, are weights that denote the inverse of the probability that the observation is included due to the sampling design.

As mentioned above, since the Spanish regions spend different amounts on students and since also the test scores vary by region, in this part is searched for a possible connection between test score differences and regional spending difference. For example -as mentioned above- the PISA results (for science) for Andalusia are the lowest of Spain. Can this be explained because Andalusia spends only a small amount, and if so, to what portion can it be explained? Also can be searched which portion of difference in science results can possibly be explained by student background characteristics and non-money related school characteristics. From the estimated coefficients of the country wide education production function, the counterfactual will be calculated. For example, it can be found what would be the score of an average student in e.g. Andalusia if he would have the same money-related average inputs as did students in Castile and Leon.

A first step in doing this is to get the mean values and the coefficients of the regression for all the variables included in the three categories, i.e. the MRV, the student background variables and the Non-MRV within schools. Given the probability to have multicollinearity in the data set, both significant and insignificant betas will be used.[7]

[7] Multicollinearity, the statistical phenomenon in which two or more predictor variables are highly correlated, does not reduce the predictive power or reliability of the model as a whole but it only affects calculations regarding individual predictors. That is, a regression with correlated predictors can indicate how well the entire bundle of predictors predicts the outcome variable, but it may not give a valid

First, a reference region is chosen. For each region and for each variable, the difference is made between the mean for this variable in the reference region and for the one in the specific community (Andalusia, Aragon, Asturias, etc.). This difference is then multiplied every time with the corresponding coefficient of the regression. By summing these products of differences for all variables in each category, it can be found how much of the difference in student performance between a certain region and the reference region can be explained by the whole model and individually by the three categories of variables.

---

results about any individual predictor, or about which predictors are redundant with others.

4. RESULTS

Using the data and the data analysis techniques described above, in this section the obtained results are discussed. In the first part of this section, the empirical results on the relationship between all the variables included in the analysis and student achievement are reported and discussed. In the second part of this section, the connection between test scores and regional spending is studied for the Spanish case.

4.1. General results

Table 4 shows the results of the country-wide regression when adding all the control variables discussed above. As shown in this table, most of the coefficients are not significant on a 5% significance level.

From the coefficients belonging to the MRV, it can be concluded that none of them is significant on a 5% significance level. Only the variable corresponding with the percentage of teachers with an ISCED5A level is significant on a 10% significance level and gives the unexpected result that a higher percentage of qualified teachers is in general associated to lower science test scores for students of that school. However, we have to keep in mind that this variable, as mentioned above, is in the best case a good approximation for what we want to measure and that we are working on only a significance level of 10%.

In the category of the NMRV, and in more detail in the subcategory of the student background variables, it can be concluded that the educational performance of students is largely associated by the grade level in which they are enrolled. Results show that students of higher grades perform on average better than those in lower grades.

In science, boys outperform girls by around 22 achievement points (AP). Given that the test scores are scaled to have an international standard deviation among the OECD countries of 100 and a standard deviation among Spain of around 90, the size of this effect can be interpreted as percentage points of a standard deviation. That is, boys perform around one quarter of a standard deviation better than girls.

From the family background coefficients it can be seen that native students perform around one third of a standard deviation better than first generation students. Students with both parents having a low education do 12,65 AP worse than students that have a parent with a high education. Students with at least one parent with a white collar and a high skill classification do better than students with a white collar and low classification as highest classification of the parents. On their turn, they do better than students with a blue collar and low classification as highest classification of the parents. Another indicator of family background that is strongly and statistically significantly related to student performance is the number of books in the students' home. High performances of students are in general associated with a high number of books in their home, leading to a difference of around a half of a standard deviation between students with more than 100 books at home and students with less than 26 books at home.

It is also important to mention that the time spent per week in science seems to have an influence on the students' performance. Both the extreme cases of devoting a little or a lot of time in science are negative associated with science test scores.

Since the remaining coefficients, including the non-money related school variables, are not significant, nothing meaningful can be said about the influence of these variables on students' achievement. From these results, it could be concluded that the non-money-related school variables used in this analysis are not associated with test scores. However, from the literature it could be expected that institutions are important for student achievement. The reason for the meaningless results for non-money related school variables is possibly due to omission of variables that should be included in the model. In this analysis a private school is defined as a school managed by a non-government organisation. However, a variable for school financing is for example not included. Another group of institutional variables that could be included in this analysis are the variables that describe the school autonomy.

By using Table 5, sums can be made of the products of the coefficients of each variable of the country-wide regression with the mean value of that variable for a certain region. From that, a comparison can be made of the Pisa test scores and estimates for all communities. From Figure 3, it can be seen that the predictions of the science scores are in general underestimations of the real PISA results. The country-wide regression gives good predictions for the test scores of Andalusia, while Castile and Leon and La Rioja are not predicted to be the best scoring regions.

4.2. Connection between test scores and regional spending

The differences between the average science score for Castile and Leon (the reference region) and the average science scores for all the other regions are calculated by using PISA results on the one hand and the results estimated by the model on the other hand (see Table 5). Figure 4 shows the results while Table 6 gives a summary of the results that are explained by the model. From this, it can be summarized that for 5 regions, i.e. Andalusia, Cantabria, Galicia, La Rioja and the rest of Spain, more than one third of the gap in test scores between Castile and Leon and that specific region can be explained. For the other half of the regions, i.e. Aragon, Asturias, Catalonia, Navarre and Basque Country, the model can hardly explain any of the gap. For these regions it is not clear why Castile and Leon outperforms these regions. This is especially true for Aragon, Catalonia, Navarre and Basque Country where these regions should be outperforming Castile and Leon, according the analysis. It can also be concluded from this analysis that in general money related variables do not have a significant effect on the difference in score between Castile and Leon and the other regions, while the major part of the explained difference between Castile and Leon and the other regions is covered by student background characteristics.

In what is following, a description is made by using Table 6 of what is predicted to be the test score in each region if students and schools in that region had the same characteristics as those in Castile and Leon. It is tried to explain why Castile and Leon has higher test scores than any other region in Spain and it is studied if this is because Castile and Leon spends more on education than do the other regions.

In Andalusia, we find from the PISA results that students score on average 45 points below Castile and Leon. Our analysis predicts that we can attribute 0,004 of these points to differences in MRV, around 23 to differences in student background characteristics, while school non-MRV attribute around 0,11 in favour of Andalusia. We are thus able to explain 23 points, i.e. around 51%. In summary, test scores for Andalusia are predicted to be almost the same if students of Andalusia would have the same money related variables than those in Castile and Leon, while test scores are expected to be on average 23 points higher if students of Andalusia would have the background of those in Castile and Leon.

Cantabria does on average around 13 points worse than Castile and Leon. However, only around 50% of this difference is explained by the model. The model predicts that if students of Cantabria would have the same background as students of Castile and Leon, they would do around 6,6 points better. However, with the same money related variables as in Castile and Leon, they would do around 0,3 points worse. From this it can be concluded that there isn't evidence that money explain this test score gap.

Galicia does around 14 points worse than Castile and Leon. The model predicts that this region does around 9,5 points worse, almost all of this difference is explained by the student background variables. For this region, around 33,5% of the difference with Castile and Leon remains unexplained.

La Rioja has almost the same PISA result for science as Castile and Leon. However, the regression overestimates the difference with around 80%, due to the prediction that students of La Rioja with the background characteristics of those in Castile and Leon would do around 2,7 points better.

Around 35% of the difference in science test score between Castile and Leon and the rest of Spain was forecasted. The positive estimated difference is completely due to the student background characteristic. Test scores for the rest of Spain are expected to be on average 13 points higher if students of the rest of Spain would have the background of those in Castile and Leon, while the same money related variables as in Castile and Leon is predicted to have a small negative impact on test scores.

As mentioned above, for the rest of the regions, the model can hardly explain any of the gap.

Aragon scores on average only around 5 points less than Castile and Leon. However, the regression model, and all the individual categories of variables, forecasts slightly better results for Aragon in the three categories. The model makes an error of around 6 points.

Instead of the real average difference of 10 points in PISA test scores between Castile and Leon and Asturias, the model only predicts a difference of around 0,35 between these regions, i.e. only around 3% can be explained. The three categories of variables only attribute for small differences, with student background and school non-MRV attributing for positive differences, while the MRV attribute for negative differences. The model predicts thus that if students from Asturias would have the same money related variables than those in Castile and Leon, they would do slightly worse, while the same having the same non-money related variables than in Castile and Leon would have a slightly positive effect on the results for Asturias.

The model predicts that Catalonia outperforms Castile and Leon with around 6,5 points. However, the reality is different, with an average score in Castile and Leon that is around 25 points higher than in Catalonia. The category of variables that is for a great part responsible for this error is the one of the student background characteristics, which attributes around 6,7 points in favour of Catalonia. Only the MRV forecast better results for Castile and Leon.

Even though Navarre is doing worse in science than Castile and Leon, the model and mainly the student background characteristics predict better results for Navarre.

The same conclusions are true for the Basque Country. For this region, the model predicts a result far away from the reality, i.e. it predicts a score for the Basque

Country that is 10 points higher than Castile and Leon while Castile and Leon is 25 points better.

Instead of taking Castile and Leon as the reference region, also Andalusia, the region with the lowest science scores, could be chosen as reference region. Table 7 gives a summary of the results that are explained by the model, when taking Andalusia as reference region.

From this, it can be summarized that for all the regions, more than one third of the gap in test scores between that specific region and Andalusia can be explained. Also from this analysis, it can be concluded that in general money related variables do not have a significant effect on the difference in score between the reference region and the other regions, while the major part of the explained difference is covered by student background characteristics.

5. SUMMARY AND CONCLUSIONS

Since the regions in Spain spend much different amounts on students and since test scores vary by region, in this thesis it is tried to find if there is a connection between test score differences and regional spending differences.

The dataset used in this analysis come from the results of 2006 of the Programme for International Student Assessment (PISA), an internationally standardised assessment that was jointly developed by 57 participating countries and administered to 15-year-olds in schools. The dataset combines average student test scores in science with student background characteristics, teacher characteristics, classroom characteristics and school characteristics.

These variables can be used in a country wide education production function for Spain. The average science test score is chosen as dependent variable, while the independent variables can be categorized in money-related-variables (MRV) and non money-related-variables (NMRV). The NMRV are subcategorized in student background variables and non-money-related school variables.

In a next step, the counterfactual was calculated from the estimated coefficients and the mean values for the regions of the variables included in the country wide education production function. First, a reference region was chosen. Predictions were made of what would be the test score in each region if students and schools in that region had the same characteristics as those in the reference region. It was tried to explain why the test has higher or lower test scores than any other region in Spain and if this is because that reference region spends more or less on education than do the other regions.

From the obtained results in this analysis for Castile and Leon as reference region, it can be summarized that for 5 regions, more than one third of the gap in test scores between Castile and Leon and that specific region can be explained. For the other part of the regions, the model can hardly explain any of the gap and most of these regions should be even outperforming Castile and Leon, according the analysis. For these regions it is thus not clear why Castile and Leon outperforms these regions.

Another conclusion from this analysis is that in general money related variables do not have a significant effect on the difference in score between Castile and Leon and the other regions, while the major part of the explained difference between Castile and Leon and the other regions is covered by student background characteristics.

Since only for the half of the regions a good share, i.e. more than one third, of the gap can be explained, there is still a big part of the difference that remains unexplained. Institutional variation in Spain might explain a part of the still unexplained gap. However, in this thesis it was not included since even in studies comparing international test scores these variables are not always significant. Is was assumed that schools in Spain would have more similarities in institutional structure than different schools on European level.

Next, the region with the lowest science scores, i.e. Andalusia, was chosen as reference region. From this analysis, it could be summarized that for all the regions, more than one third of the gap in test scores between that specific region and Andalusia could be explained. From this, it can be concluded that the model performs different by explaining test score differences between regions in Spain when choosing another reference region. However, also in this analysis it was concluded that in general money related variables do not have a significant effect on the difference in score between the reference region and the other regions and that the major part of the explained difference is covered by student background characteristics.

In summary, from this paper it is suggested that educational policy in Spain should not focus on providing more resources to schools. It seems that differences in test scores are explained at least for a part by student background characteristics. However, there still needs to be explained a significant part of the difference in test scores between regions in Spain. Future research may include institutional variation in Spain to explain a part of this still unexplained gap. Another limitation of this study for explaining test score differences by different groups of variables is the probability to have multicollinearity in the used data set, since some predictor variables may be highly correlated. Therefore, the regression may not give valid results about any individual predictor or group of predictors.

6. Tables

COMMUNITY	SCIENCE SCORE	READING SCORE	MATHEMATICS SCORE
Castile and Leon	518,43	476,7	513,85
La Rioja	517,33	488,62	523,34
Navarre	513,67	482,48	518,5
Aragon	513,39	484,84	513,37
Asturias	508,18	474,67	494,74
Cantabria	505,62	470,6	494,17
Galicia	504,12	475,62	491,1
Basque Country	493,5	488,54	500,43
Catalonia	493,11	476,26	491,66
Spain (country)	488	460	479
Rest of Spain*	482,31	450	470,37
Andalusia	*473,72*	*444,68*	*461,93*

Table 1. Mean of scores obtained by each community in PISA 2006; the communities are shown according their ranking in science score. *The rest of Spain refers to: Canarias, Castile-La Mancha, Extremadura, Illes Balears and Murcia. Madrid and Valencia did not present the examination. Source: Table made with PISA 2006 database.

COMMUNITY	EDUCATIONAL EXPENDITURE	REGISTERED STUDENT POPULATION	PER PUPIL EXPENDITURE
Basque Country	772.506.700	100.218	7.708
Navarra	188.175.660	29.013	6.486
Asturias	278.157.800	45.021	6.178
Castile and Leon	671.786.000	124.057	5.415
La Rioja	75.738.890	14.507	5.221
Catalonia	1.758.970.782	341.149	5.156
Spain (country)	**11.462.952.651**	**2.334.327**	**4.911**
Aragon	303.781.640	62.695	4.845
Rest of Spain	4.621.377.283	959.724	4.815
Galicia	640.945.676	136.065	4.711
Andalusia	2.040.183.000	494.698	4.124
Cantabria	111.329.220	27.179	4.096

Table 2. Educational Expenditure (in miles of Euros) and registered student E.S.O. and F. Professional population. The communities are shown according their per pupil expenditure. Source: Table made with Education Investment Statistics (MEC, 2007) and Education Data and Numbers (MEC, 2007).

	VARIABLE	Mean	Std. Dev.	Obs.
	Science test score	487,90	89,76	23.266
MRV	**TEACHER CHARACTERISTICS**			
	Teachers' certification			
	Percentage of teachers that are fully certified	0,818	0,313	13.222
	Percentage of teachers with an ISCED5A classification	0,718	0,361	14.573
	Lack of qualified science teachers			
	Not at all-very little	0,787	0,409	23.266
	To some extent-a lot	0,159	0,365	23.266
	CLASSROOM CHARACTERISTICS			
	Average size of class			
	15 students or less	0,063	0,243	23.266
	Between 16 and 30	0,592	0,491	23.266
	Between 31 and 50	0,232	0,422	23.266
	51 students or more	0,066	0,248	23.266
	SCHOOL CHARACTERISTICS			
	Computers available at school			
	Ratio PC student	0,165	0,172	21.682
	Shortage of material			
	Not at all-very little	0,255	0,436	23.266
	To some extent-a lot	0,096	0,294	23.266
NMRV	**STUDENT BACKGROUND CHARACTERISTICS**			
	Grade			
	7^{th}	0,001	0,027	23.266
	8^{th}	0,077	0,267	23.266
	9^{th}	0,336	0,472	23.266
	10^{th}	0,586	0,493	23.266
	11^{th}	0,000	0,017	23.266
	Gender			
	Male	0,511	0,500	23.266
	Female	0,489	0,500	23.266
	Immigration status			
	Native	0,917	0,275	23.266
	First generation	0,063	0,243	23.266
	Second Generation	0,007	0,082	23.266
	Highest educational level of parents			
	Low (ISCED 0,1 or 2)	0,367	0,482	23.266
	Medium (ISCED 3 or 4)	0,272	0,445	23.266
	High (ISCED 5 or 6)	0,344	0,475	23.266
	Highest white/blue collar classification of parents			
	White collar high skilled	0,372	0,483	23.266
	White collar low skilled	0,259	0,438	23.266
	Blue collar high skilled	0,238	0,426	23.266
	Blue collar low skilled	0,107	0,309	23.266
	Books at home			
	0-25 books	0,224	0,417	23.266
	26-100 books	0,329	0,470	23.266
	More than 100 books	0,431	0,495	23.266
	Time spent per week in school science			
	Little Time (<2 hours per week in regular lessons, out-of-school-	0,216	0,411	23.266
	Much Time(>2 hours per week in regular lessons, out-of-school-	0,677	0,251	23.266
	SCHOOL CHARACTERISTICS			
	Public vs. Private school			
	Public school	0,783	0,412	23.266
	Private school	0,164	0,370	23.266
	Community in which school is located			
	Village, rural area or town (less than 100000 people)	0,643	0,479	23.266
	City or Large City	0,317	0,465	23.266

TABLE 3. DESCRIPTIVE STATISTICS OF THE VARIABLES USED. NOTE: OBSERVATIONS ARE WEIGHTED BY STUDENT WEIGHTS.

	VARIABLE	Coef.		Std. Error
MRV	**TEACHER CHARACTERISTICS**			
	Teachers' certification			
	Percentage of teachers that are fully certified	1,858		3,707
	Percentage of teachers with an ISCED5A classification	-4,893	*	2,893
	Lack of qualified science teachers			
	Not at all-very little	REFERENCE[1]		
	To some extent-a lot	1,067		2,722
	CLASSROOM CHARACTERISTICS			
	Average size of class			
	15 students or less	REFERENCE[1]		
	Between 16 and 30	3,710		4,084
	Between 31 and 50	2,543		4,693
	51 students or more	-7,423		6,022
	SCHOOL CHARACTERISTICS			
	Computers available at school			
	Ratio PC student	3,104		6,232
	Shortage of material			
	Not at all-very little	-2,712		2,439
	To some extent-a lot	4,564		4,088
	Else	REFERENCE[1]		
NMRV	**STUDENT BACKGROUND CHARACTERISTICS**			
	Grade			
	7^{th}	-175,360	***	11,667
	8^{th}	-113,578	***	4,282
	9^{th}	-70,274	***	2,530
	10^{th}	REFERENCE[1]		
	11^{th}	176,074	***	51,074
	Gender			
	Male	REFERENCE[1]		
	Female	-21,673	***	2,137
	Immigration status			
	Native	32,542	***	5,577
	First generation	REFERENCE[1]		
	Second Generation	29,859	**	12,157
	Highest educational level of parents			
	Low (ISCED 0,1 or 2)	-12,655	***	2,940
	Medium (ISCED 3 or 4)	-4,810	*	2,821
	High (ISCED 5 or 6)	REFERENCE[1]		
	Highest white/blue collar classification of parents			
	White collar high skilled	24,956	***	4,041
	White collar low skilled	10,739	***	3,825
	Blue collar high skilled	9,002	**	3,782
	Blue collar low skilled	REFERENCE[1]		
	Books at home			
	0-25 books	REFERENCE[1]		
	26-100 books	26,642	***	3,101
	More than 100 books	45,735	***	3,251
	Time spent per week in school science			
	Little Time (<2 hours per week in regular lessons, out-of-school-time lessons and study or homework)	-31,691	***	2,582
	Much Time(>2 hours per week in regular lessons, out-of-school-time lessons and study or homework)	-38,170	***	4,062
	Else	REFERENCE[1]		
	SCHOOL CHARACTERISTICS			
	Public vs. Private school			
	Public school	-2,094		2,757
	Private school	REFERENCE[1]		
	Community in which school is located			
	Village, rural area or town (less than 100000 people)	REFERENCE[1]		
	City or Large City	-0,428		2,390

TABLE 4. RESULTS OF THE REGRESSION. NOTE: ***, **,* REPRESENT SIGNIFICANCE AT 99%, 95% AND 90% RESPECTIVELY.

[1] THE REFERENCE GROUP ALSO CONTAINS A SMALL NUMBER OF MISSING VALUES.

VARIABLE	Regression Coefficients	SPAIN	Andalusia	Aragon	Asturias	Cantabria	Castile and Leon	Catalonia	Galicia	La Rioja	Navarre	Basque Country	Rest of Spain
Science test score		487,90	473,72	513,39	508,18	505,62	518,43	493,11	504,12	517,33	513,67	493,50	482,31
Fully Certified teachers (%)	1,86	0,82	0,85	0,84	0,84	0,88	0,84	0,86	0,80	0,88	0,86	0,84	0,78
ISCED5A certified teachers (%)	-4,89	0,72	0,70	0,73	0,75	0,72	0,75	0,74	0,78	0,71	0,72	0,74	0,70
Shortage of qualified science teachers	1,07	0,16	0,14	0,17	0,18	0,19	0,24	0,13	0,16	0,24	0,17	0,18	0,16
Class size 16 - 30 students	3,71	0,59	0,58	0,63	0,64	0,58	0,50	0,59	0,57	0,63	0,52	0,57	0,62
Class size 31 - 50 students	2,54	0,23	0,23	0,22	0,24	0,25	0,32	0,23	0,27	0,19	0,26	0,25	0,21
Class size 51 and + students	-7,42	0,07	0,09	0,04	0,06	0,08	0,10	0,10	0,07	0,03	0,08	0,06	0,04
Ratio PC student	3,10	0,16	0,15	0,16	0,15	0,17	0,14	0,16	0,15	0,19	0,15	0,16	0,18
No shortage of material	-2,71	0,25	0,25	0,35	0,21	0,23	0,22	0,25	0,23	0,26	0,19	0,28	0,26
Shortage of material	4,56	0,10	0,08	0,12	0,10	0,09	0,13	0,09	0,09	0,08	0,13	0,09	0,10
7 th grade	-175,36	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
8 th grade	-113,58	0,08	0,11	0,08	0,07	0,06	0,05	0,03	0,13	0,05	0,04	0,05	0,08
9 th grade	-70,27	0,34	0,39	0,29	0,27	0,37	0,35	0,30	0,28	0,36	0,25	0,21	0,35
11 th grade	176,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Female	-21,67	0,49	0,47	0,49	0,48	0,43	0,45	0,48	0,45	0,44	0,52	0,52	0,51
Native	32,54	0,92	0,97	0,92	0,95	0,94	0,96	0,88	0,96	0,89	0,90	0,95	0,88
Second Generation	29,86	0,01	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	0,01
Parent Low Educated (ISCED 0 - 2)	-12,65	0,37	0,48	0,29	0,20	0,26	0,29	0,30	0,36	0,34	0,24	0,22	0,38
Parent Medium Educated (ISCED 3 -4)	-4,81	0,27	0,27	0,29	0,36	0,33	0,29	0,25	0,33	0,31	0,29	0,30	0,26
White collar high skilled	24,96	0,37	0,28	0,42	0,39	0,40	0,34	0,47	0,34	0,39	0,39	0,41	0,38
White collar low skilled	10,74	0,26	0,27	0,28	0,28	0,23	0,29	0,24	0,23	0,25	0,30	0,26	0,26
Blue collar high skilled	9,00	0,24	0,34	0,17	0,20	0,17	0,23	0,19	0,31	0,19	0,18	0,18	0,21
Books at home 26 -100	26,64	0,33	0,40	0,32	0,33	0,33	0,33	0,32	0,32	0,29	0,34	0,31	0,30
Books at home 101 or +	45,74	0,43	0,30	0,52	0,47	0,47	0,53	0,46	0,46	0,53	0,51	0,51	0,45
Little time spent on science	-31,69	0,22	0,24	0,19	0,24	0,22	0,20	0,09	0,26	0,17	0,28	0,21	0,24
Much time spent on science	-38,17	0,07	0,06	0,05	0,10	0,08	0,06	0,04	0,06	0,07	0,04	0,04	0,09
Public school	-2,09	0,78	0,76	0,78	0,81	0,83	0,81	0,80	0,72	0,80	0,82	0,77	0,79
City	-0,43	0,32	0,31	0,31	0,41	0,35	0,31	0,36	0,33	0,30	0,33	0,30	0,30

Table 5. Mean values and the coefficients of the regression for all the variables included in the three categories for all communities. The constant value of the regression is 475.
Source: Elaborated with the regression data.

GAP with respect to Castile and Leon		Andalusia	Aragon	Asturias	Cantabria	Catalonia	Galicia	La Rioja	Navarre	Basque Country	Rest of Spain
PISA Gap		**44,71**	**5,04**	**10,25**	**12,81**	**25,32**	**14,31**	**1,09**	**4,75**	**24,93**	**36,12**
Explained by the model:											
MRV		0,00	-0,34	-0,46	-0,32	0,16	0,13	-0,74	-0,28	-0,08	-0,50
NMRV	Student Background	23,08	-0,07	0,76	6,64	-6,67	9,56	2,74	-3,85	-9,82	13,09
	School	-0,11	-0,07	0,04	0,05	0,00	-0,18	-0,03	0,01	-0,10	-0,06
Total Explained		**22,97**	**-0,48**	**0,35**	**6,37**	**-6,51**	**9,51**	**1,97**	**-4,11**	**-10,01**	**12,53**
Unexplained		21,74	5,51	9,90	6,45	31,83	4,80	0,88	8,87	34,93	23,58

Table 6. Differences between real PISA 2006 gap and model explanations with Castile and Leon as reference region. Source: Elaborated with PISA 2006 database and model regression.

GAP with respect to Andalusia		Aragon	Asturias	Cantabria	Castile and Leon	Catalonia	Galicia	La Rioja	Navarre	Basque Country	Rest of Spain
PISA Gap		**-39,67**	**-34,46**	**-31,90**	**-44,71**	**-19,40**	**-30,40**	**-43,62**	**-39,96**	**-19,79**	**-8,59**
Explained by the model:											
MRV		-0,35	-0,46	-0,33	0,00	0,15	0,12	-0,74	-0,28	-0,09	-0,50
NMRV	Student Background	-23,15	-22,32	-16,44	-23,08	-29,75	-13,52	-20,34	-26,93	-32,90	-9,99
	School	0,05	0,16	0,17	0,11	0,11	-0,07	0,09	0,13	0,01	0,06
Total Explained		**-23,45**	**-22,62**	**-16,60**	**-22,97**	**-29,48**	**-13,46**	**-21,00**	**-27,09**	**-32,98**	**-10,44**
Unexplained		-16,23	-11,84	-15,30	-21,74	10,08	-16,94	-22,62	-12,87	13,19	1,84

Table 7. Differences between real PISA 2006 gap and model explanations with Andalusia as reference region. Source: Elaborated with PISA 2006 database and model regression.

7. FIGURES

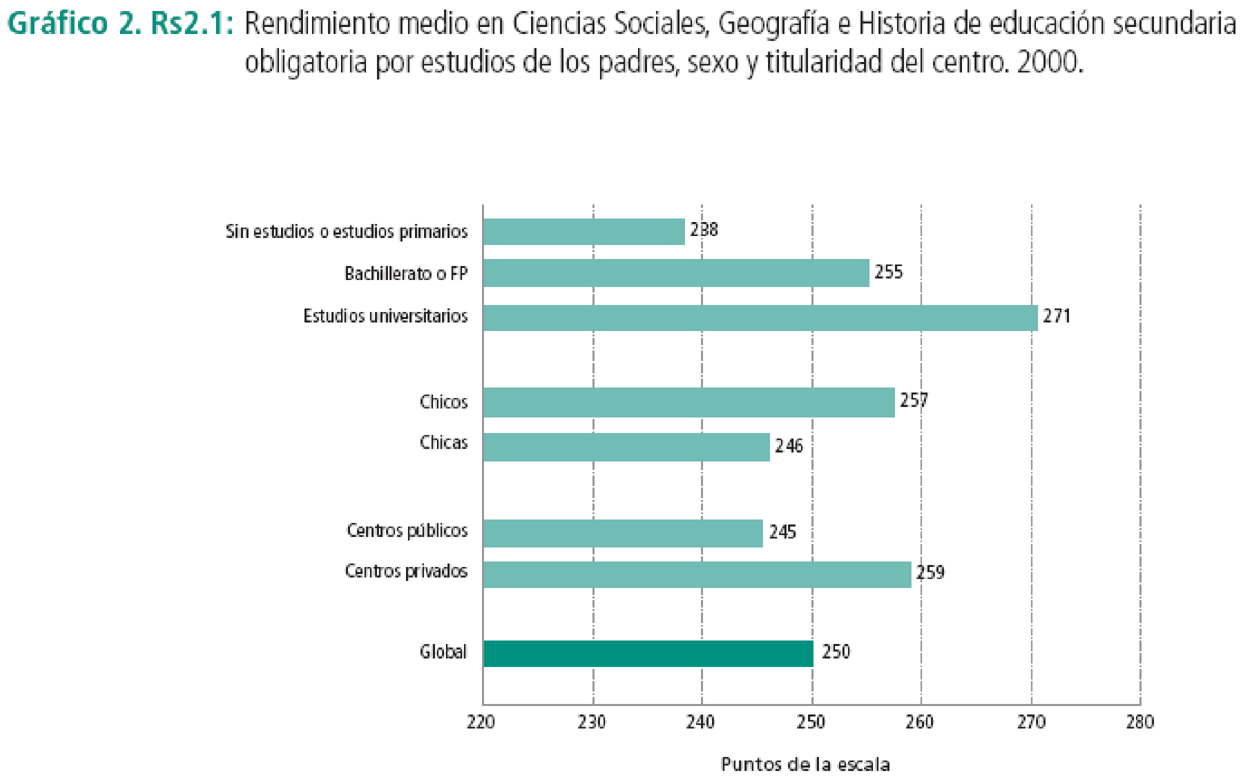

FIGURE 1. MEAN PERFORMANCE OF E.S.O. STUDENTS IN SCIENCE, GEOGRAPHY AND HISTORY ACCORDING TO THEIR GENDER, SCHOOL OWNERSHIP AND THEIR PARENTS EDUCATION. SOURCE: HIGH-PRIORITY INDICATOR REPORT FROM M.E.C. (MARTÍN, 2007).

Gráfico 7.E5: Alumnado extranjero por cada mil alumnos en cada etapa educativa según Comunidad Autónoma. Cursos 2002-03 y 2003-04.

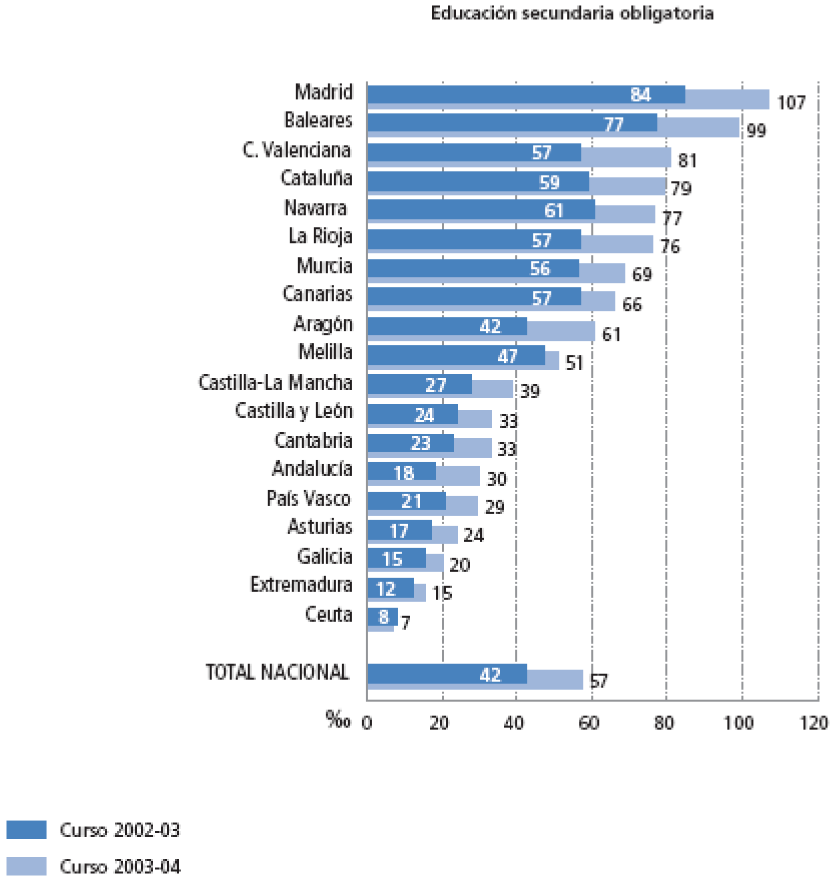

FIGURE 2. REGISTERED FOREIGN E.S.O. STUDENT POPULATION IN EACH OF THE AUTONOMOUS COMMUNITIES. SOURCE: HIGH-PRIORITY INDICATOR REPORT FROM M.E.C. (MARTÍN, 2007).

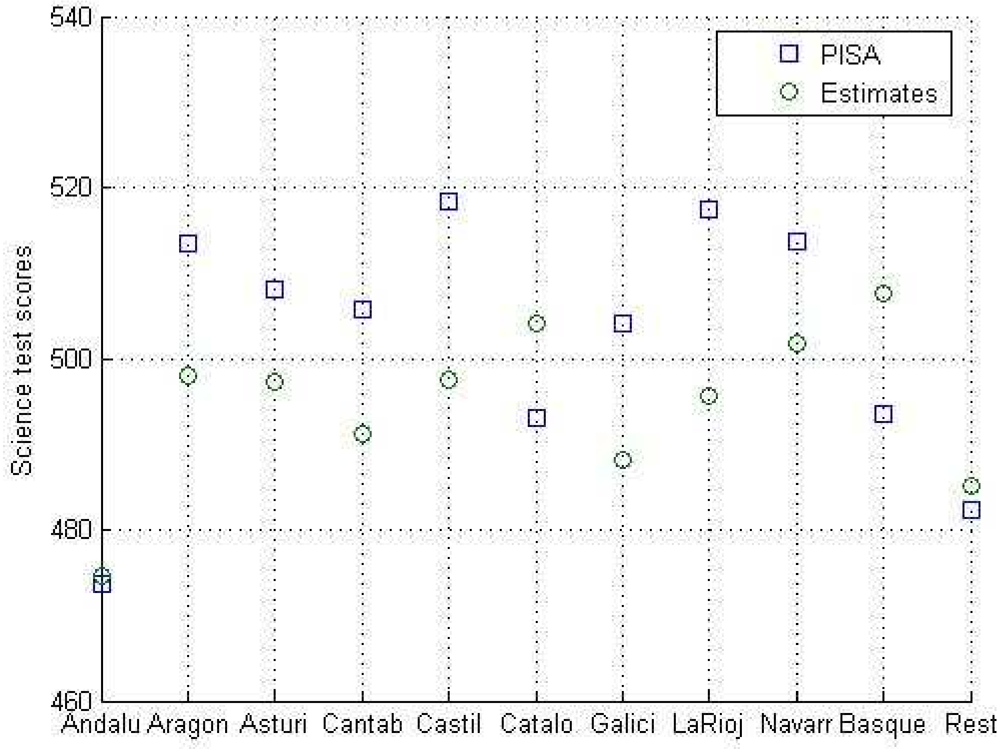

FIGURE 3. PISA TEST SCORES AND ESTIMATES FOR ALL COMMUNITIES. SOURCE: ELABORATED WITH REGRESSION INFORMATION.

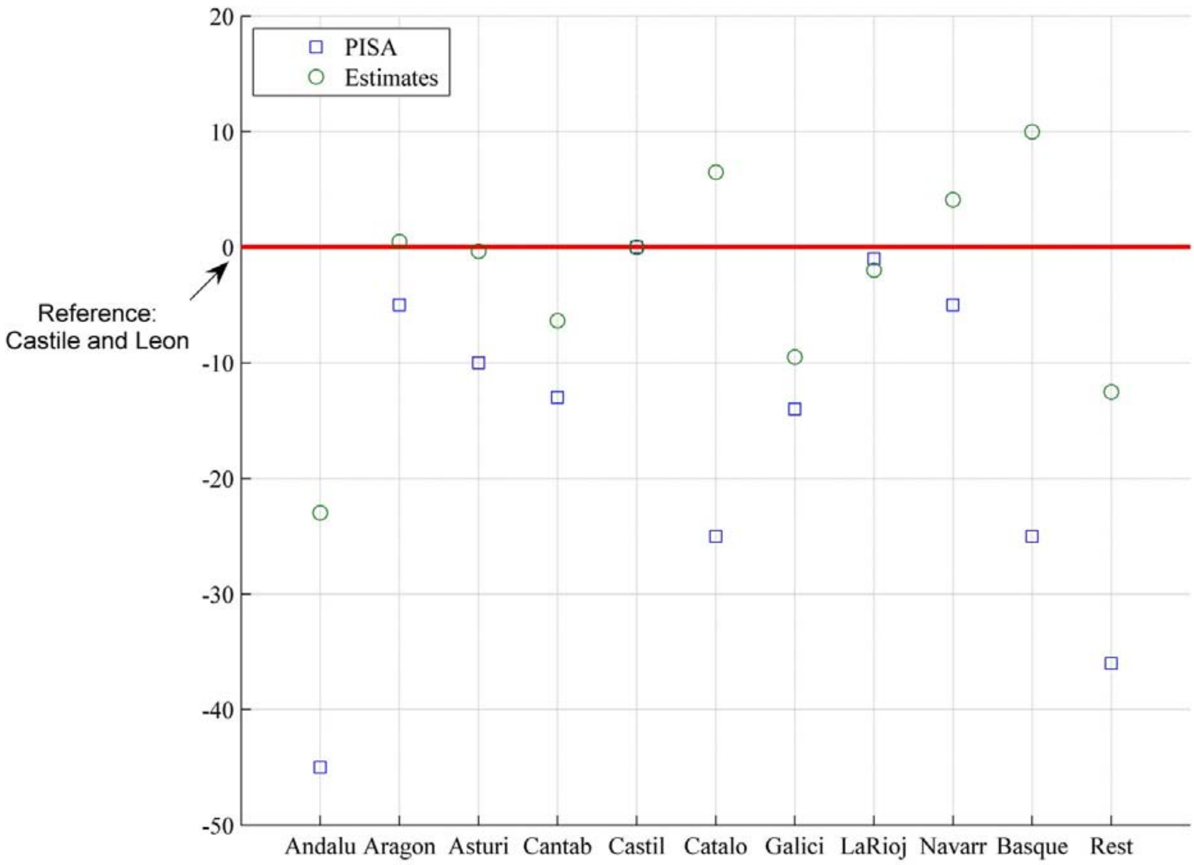

FIGURE 4. DIFFERENCES IN SCIENCE SCORES COMPARED TO CASTILE AND LEON. SOURCE: MADE WITH PISA 2006 DATABASE AND REGRESSION RESULTS OF THIS THESIS.

8. References

1. Agencies. (2007). "La inversión en educación en España sigue siendo baja respecto a los países de la OCDE." El Mundo. Spain: Mundinteractivos, S.A. Consultedon May 2008. <http://www.elmundo.es/elmundo/2007/10/02/espana/1191337668.html>
2. Agencies. "España 'suspende' en Educación y Andalucía saca las peores notas". (2007). El Mundo. Spain: Mundinteractivos, S.A. Consulted: May 2008. <http://www.elmundo.es/elmundo/2007/12/04/espana/1196763523.html>
3. Datos Básicos de la educación en España en el curso 2007. Spain: Ministerio de Educación y Ciencia; Subdirección General de Información y Publicaciones. 2007. Consulted: July 2008. <http://www.mepsyd.es/mecd/estadisticas/educativas/dcce/DATOS_Y_CIFRAS_WEB.pdf>
4. Entorf, Horst and L. Martina. (2007). "Peer Effects, Social Multipliers and Migrants at School". CEGE Discussion Paper 57. Germany: Georg – August-Universität Götingen.
5. "Estadística del Gasto Público en Educación". (2007)Spain: Ministerio de Educación, Política Social y Deporte. Consulted: June 2008. <http://www.mepsyd.es/mecd/jsp/plantilla.jsp?id=311&area=estadisticas&contenido=/estadisticas/educativas/gasto/2005/gastos2005.html>
6. Fertig, Michael. (2004) "What can we learn from international student performance studies?" Germany: Rheinisch-Westfälisches Institut für Wirtschaftsforschung. Consulted: June 2008.
7. FETE – UGT. (2007). Informe PISA 2006. Spain: Gabinete Técnico FETE.
8. Fuchs, Thomas and Wöbmann. (2004a) "Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and use of computers at home and at school". Germany: Ifo Institute for Economic Research at the University of Munich. Consulted: May 2008.
9. Fuchs, Thomas and Woessmann. (2004b). "What Accounts for International Differences in Student Performance?, A Re-Examination Using PISA Data." Munich: Ifo Institute for Economic Research at the University of Munich.
10. Gamoran, Adam and Long. (2006). "Equality of Educational Opportunity: A 40-year retrospective". Madison: University of Wisconsin-Madison. Consulted: June 2008 . <http://www.wcer.wisc.edu/publications/workingPapers/papers.php>
11. Goenechea, Cristina. (2006). "Diagnóstico de la situación educativa de los alumnos extranjeros en España". Spain: Universidad Complutense de Madrid. Consulted: July 2008.<http://weib.caib.es/Documentacio/jornades/Web_I_Cong_Medit/PDFs/diagnostic2.pdf>
12. Gundlach, Erich and L. Wöbmann. (2001). "Better Schools for Europe". Germany: EIB paper, volumen 6 n.2, 2001.
13. Instituto de Evaluación en las Comunidades Autónomas. (2008). Spain: Ministerio de Educación, Política Social y Deporte. Consulted: June 2008. <http://www.institutodeevaluacion.mec.es/comunidades_autonomas/>
14. Junta de Andalucía.(2007). "Informe sobre el sistema educativo de Andalucía (2002-2007)". Spain: Consejería de Educación de Andalucía. Consulted: June 2008. <http://www.juntadeandalucia.es/educacion/scripts/w_cea/informes.htm>

15. Lacasa, José M. (2008) "Hacia un desequilibrio regional originado por el fracaso escolar en la ESO". Spain: Edición Digital de Magisterio. Consulted: June 2008. <http://www.magisnet.com/articulos.asp?idarticulo=2251&n_edicion=11728>
16. Martín, Joaquín, Ayllón, Benedí, Bonilla. (2007). "Sistema estatal de indicadores de la educación prioritarios 2006". Madrid: Ministerio de Ciencia y Educación, Instituto de Evaluación, 2007. Consulted: June 2008.<http://www.institutodeevaluacion.mec.es/contenidos/indicadores/prioritarios2006CD.pdf>
17. Mora, Toni, Escardíbul and Espasa.(2006). "The effects of regional education policies on school failure rates in Spain". Catalunya: Universitat Internacional de Catalunya. Consulted: June 2008. <http://sadapt.inapg.inra.fr/ersa2007/papers/17.pdf>
18. Ombudsman – UNICEF.(2007). "Violencia Escolar: El maltrato entre iguales en la educación secundaria obligatoria 1999-2006. Nuevo estudio y actualización del informe 2000". Madrid: Defensor del Pueblo, 2007. Consulted: August 2008.<http://www.defensordelpueblo.es/informes2.asp>
19. PISA 2006.(2007). Centros Educativos de Aragón. Educaragón: Gobierno de Aragón. 2007. Consulted: June 2008. <http://www.educaragon.org/arboles/arbol.asp?guiaeducativa=41&strseccion=A1A94>
20. PISA 2006. (2007). Database download. Ed: OECD. Consulted: June 2008. <http://pisa2006.acer.edu.au/downloads.php>
21. Pritchett, Lant and Filmer.(1997). "What educational production functions really show, a positive theory of education spending." The World Bank. Consulted: June 2008.
22. Schleicher, Andreas.(2006)."OECD Briefing Note for Spain". Consulted: May 2008. <http://www.oecd.org/dataoecd/22/31/39317551.pdf>
23. Swedberg-Gonzalez, Pablo.(2007). "Spanish students rank low in the sciences". Spain: 5 Spaniards 3 December 2007. Consulted: May 2008.<http://5spaniards.com/en/index.php?option=com_content&task=view&id=59&Itemid=102>
24. USTEA.(2008). "Retribuciones a profesores 2008". Andalusia: Unión de Sindicatos de Trabajadores y Trabajadoras en Andalusia. 2008. Consulted: May 2008. <http://www.ustea.org/educacion/>