6-1-2015

Non-Cognitive Abilities and Spanish Regional Differences in Student Performance in PISA 2009

Ildefonso Mendez
University of Murcia

Gema Zamarro
University of Arkansas, Fayetteville, gzamarro@uark.edu

Jose G. Clavel
University of Murcia

Collin Hitt
University of Arkansas

Citation

Mendez, I., Zamarro, G., Clavel, J. G., & Hitt, C. (2015). Non-Cognitive Abilities and Spanish Regional Differences in Student Performance in PISA 2009. *Education Reform Faculty and Graduate Students Publications.* Retrieved from https://scholarworks.uark.edu/edrepub/44

char·asse·in

THE CHARACTER ASSESSMENT INITIATIVE

NON-COGNITIVE ABILITIES AND SPANISH REGIONAL DIFFERENCES IN STUDENT PERFORMANCE IN PISA 2009

Ildefonso Méndez (Universidad of Murcia)
Gema Zamarro (University of Arkansas & University of Southern California)
José G. Clavel (Universidad of Murcia)
Collin Hitt (University of Arkansas)

[ORIGINAL ARTICLE IN SPANISH WRITTEN FOR THE JOURNAL "Participación Educativa", PUBLISHED BY THE SPANISH "CONSEJO ESCOLAR DEL ESTADO", SPANISH MINISTRY OF EDUCATION, CULTURE AND SPORTS]

[First Version: March, 2015. English Translation Version: June, 2015]

EDRE Working Paper No. 2015-05
Last Updated June 2015

Abstract:

The goal of this paper is to analyze the role that non-cognitive skills and, in particular, regional differences in those skills, play on the observed differences in 15-year-old student's academic performance, across Spanish regions, on PISA 2009. Previous research has shown the relevance of differences in student's personal, family and school characteristics in accounting for academic differences across Spanish regions but it has also found that a sizeable part of the observed differences remained unexplained. We have found that differences in the distribution of certain non-cognitive skills associated to academic performance like focus, perseverance and resilience play a prominent role in accounting for differences in student performance in PISA 2009. We observe these skills by developing new measures of student effort on standardized tests. In particular, our estimates suggest that a standard deviation reduction in the dispersion of non-cognitive skills across Spanish regions would lead to a 25% reduction in the magnitude of the observed differences in student performance across regions. This is a relevant effect as, for example, a one standard deviation reduction in the regional dispersion of parent's educational levels or occupational status would only lead to at most a 2% reduction in the magnitude of observed differences in performance on PISA across Spanish regions. Put plainly, a substantial portion of the regional variation in test scores appears attributable to effort on the PISA test, and not necessarily just differences in actual knowledge.

Keywords: Non-cognitive skills, perseverance, resilience, academic performance, PISA study, Spain.

1. Introduction

Since 1966, when James Coleman published in his famous report that the role played by schools in shaping student's educational outcomes was almost irrelevant outside the family context, research in the area of the determinants of students' academic achievement has made great progress. There is now a better understanding of what are the main factors that contribute to better academic outcomes. Due to the advancement of econometric techniques and the increased availability of data, fifty years later we can say that Coleman's original claim is disputable. He admitted this himself in his later work: school characteristics appear to shape student outcomes.

However, the debate still persists. In fact, thanks to large scale international assessments that became to be available at the end of the last century—assessments such as PISA (Program for International Student Assessment) promoted by the Organization for Economic Cooperation and Development (OECD), TIMSS (Third International Mathematics and Science Study), and PIRSL (Progress in International Reading Literacy Study) promoted by the IEA (International Association for the Evaluation of Educational Achievement)—we have been able to better study how much of student achievement is due to each of the areas that have been traditionally thought to influence it: the characteristics of schools (e.g. school organization, educational practices, availability of resources); student's family background (e.g. socio-economic conditions, parental education, country of origin), and the circumstances in which their education develops (e.g. peer effects, place of residence, role models available).

As was recently summarized by Gil and Sanz (2014), if we study the data collected by PISA in the five waves available from PISA 2000 to PISA 2012, for the case of OECD countries, about 32% of the variation in student academic performance is due to factors attributable to the school, while another 10% would be attributable to country characteristics, and the remaining 58% would be explained by students' individual characteristics.

A literature review by Cordero, Crespo and Pedraja (2013) covered the major research contributions in this area of research as it pertains to the Spanish context, and classified the different papers according to the different methodology approaches used. In particular, Cordero, Crespo and Pedraja (2013) identify two types of studies depending on whether they use a regression analysis or an efficiency frontier approach. For both type of models, their literature review emphasized the importance of variables related to socio-economic characteristics, being held back a grade and the concentration of migrant students in school, to explain regional differences in PISA performance in Spain. Similarly, variables related to the school characteristics had a much lower predicted power. In this respect, Gonzalez and de la Rica (2012) found similar conclusions taking advantage of the increased number of regions in Spain that individually participated in PISA with a full sample in 2009. In this year specifically, only the Spanish regions of Castile-La Mancha, Valencia, and Extremadura did not participate with a full sample in the PISA study.

From this last work we would like to point out the following important result thatemerges even after controlling for the traditional factors in the model. Gonzalez and de la Rica (2012) conclude that there are significant regional disparities in student academic performance that the model is unable to explain, and whose impact is captured in the model by significant remaining regional fixed effects. This same result is also found by Villar (2012) who further quantified the differences by using several composite measures of student achievement including educational poverty, insufficient educational training, and the Educational Development Index, among others. Gonzalez and de la Rica (2012) then, proposed and tested several regional macroeconomic variables to try to explain the observed differences. These included: regional GDP, qualified employment rates, public expenditures on education, among others. But a significant portion of the regional differences still remained unexplained.

In this paper, we follow this line of research but we approach our analysis of possible determinants of the observed regional differences in student performance from a microeconomic point of view. Using the same database (PISA 2009), we try to explain regional differences by analyzing the characteristics of the smallest but most important agent in the educational process, the student, whose personal characteristics, as aforementioned, account for more than half of the variation in academic performance.

Specifically, following the proposal of James Heckman (the Nobel Prize winner in Economics in 2000; see Heckman and others, 2006), we introduce measures of non-cognitive skills of students into the analysis as a possible factor to explain regional differences in academic performance. By non-cognitive skills, we are referring to all those attitudes, behaviors, and strategies that, apart from cognitive ability, facilitate success in the work environment, and in our case in school. In particular, we are interested in character traits such as grit, perseverance and self-control. We distinguish non-cognitive skills from cognitive skills such as intelligence or standardized test scores, which might be harder to be altered specially after a very early age, but which has been the focus of most education research in recent decades.

The academic performance measures that are the focus of the previously described international assessments are mostly measures of cognitive skills. For instance the six different levels of reading ability measured in PISA 2009 reflect levels of cognitive abilities. Although some exceptions exist like in PISA 2006, where some questions related to student motivation were included in the questionnaires. These types of questions are rare and in most cases have not been asked in a stable fashion in consecutive waves. Our analysis does not rely on self-reported measures instead we argue that the PISA test can be considered as a behavioral task which provides relevant information about non-cognitive skills. Thus, by studying the patterns of response of students in the PISA 2009 test we are going to be able to infer and construct intrinsic measures of non-cognitive skills.

In recent years, personality psychologists have developed various measures of non-cognitive skills mostly based on self-reports collected by questionnaires. These measures, however, are not available in many databases, and specifically not available in PISA. In addition, these measures based on self-reports have also been found to have their limitations. This is due in large part because the validity relies on students fully responding to the questionnaire in a truthful manner, and on students having the same intrinsic standards when evaluating their skills.

In fact, measures based on the detailed study of the pattern of students’ responses to questionnaires and tests have been proposed as an alternative powerful measure of non-cognitive skills. These alternative measures capitalize on the fact that data collected in a survey contains information not only about the specific answers to the questions presented, but also about the level of commitment of demonstrated by the participants taking the assessment. For example, it is not unusual for some respondents to skip questions or to say "I don't know the answer," and it is this behavior that tells us something about the non-cognitive skills of these participants. In this sense, the work of Hitt, Trivitt and Cheng (2015) shows that the frequency with which students skip questions in a survey significantly predicts student's educational attainment and their level of income later in life, independently of the assessed level of cognitive ability. Using the fact that students participating in PISA are given different test booklets, with the questions presented in a different order, Borghans and Schils (2013) were able to quantify the rate of decline in performance in the test as the test progresses. The authors were then able to show that the rate of decline in performance over the course of the test's administration was a good predictor of final levels of educational attainment. Our work builds on this literature. We build measures of student's non-cognitive skills based on the pattern of response to the tests that were part of the PISA study in 2009. As in Borghans and Schils (2013) we also build measures of non-cognitive skills based on the rate of decline in performance in the test as the test progresses.

In a related work Balart and Cabrales (2014) analyzed Spanish regional differences in the rate of decline in performance on the test using different waves of PISA. However, unlike Balart and Cabrales (2014), we further study the possible sources of the observed uneven pace of decline in performance on PISA by analyzing additional aspects of student's patters of response to the test that have not yet been studied in the literature. For example, we study the extent to which students in different regions in Spain are more or less likely to leave questions blank and move to the next question, following the work of Hitt, Trivitt and Cheng (2015). Similarly, we also study the extent to which students are able to organize themselves to be able to respond to the whole questionnaire in the time initially established, and the extent to which they are more or less able to overcome poor initial test performance and improve as the test progresses. In this sense, we develop and study new measures of non-cognitive skills that seek to capture the level of resilience of students on tests. Resilience can be defined as the "human ability to adapt to extreme situations and overcome them." Once we build previous measures of non-cognitive skills based on students' response patterns to the PISA test, we study the degree to which they vary across the different regions in Spain, holding students' individual, family and school characteristics constant.

The rest of this paper goes as follows. Section 2 presents the Data and Empirical Model used for the analysis. Section 3 presents our results on the study of regional differences in non-cognitive skills in Spain. Finally, Sections 4 and 5 present our final discussion and conclusions.

2. Approach

2.1 Data and measures of non-cognitive skills

The PISA program sponsored by the OECD since 1997 aims to study three basic skills among 15-year-old students across countries (reading comprehension, science and mathematics skills). The program aims to measure, not what the student can do, but what the student can do with what he/she knows. As an international assessment, PISA has grown from 32 countries in 2000 to 65 in 2009, including 32 associated countries outside the OECD.

In 2009, reading comprehension was the most heavily tested basic skill assessed by the PISA program. Across Spain, the test was administered to nearly 26,000 students in 910 schools. For the first time in history, students in the Autonomous Communities (CCAA)[1] of the Canary Islands, Murcia, Madrid, and the Balearic Islands, in Spain, participated in the PISA study. These four new regions were added to the ten others that participated and provided full samples in previous waves of the PISA study. Only three of the Autonomous Communities of Spain did not participate in the PISA study, which provides us with an interesting, almost full sample of all of the regions in Spain for our analysis.

[1] There are 17 Autonomous Communities in Spain. These are: Andalusia, Aragon, Asturias, Balearic Islands, Basque Country, Canary Islands, Cantabria, Castile-La Mancha, Castile and León, Catalonia, Extremadura, Galicia, La Rioja, Madrid, Murcia, Navarre, and Valencian Community. Spanish Autonomous Communities are a first-level political and administrative division governed according to the Spanish Constitution and their own organic laws. The scope of competences vary for each community, but all have adopted competences related to the provision of education. For this paper we will often refer to autonomous communities in Spain as regions but it should be stressed out that we always mean autonomous communities. To the previously described autonomous communities we should add the cities of Ceuta and Melilla that, although they do not form their own autonomous communities, are also part of Spain and as such are included in the PISA study.

While the results of PISA 2012 were readily available at the time of this writing, and, in fact, the next wave of PISA testing took place in April 2015 in Spain, the results presented here are based on the 2009 PISA data. This is the case because the PISA 2012 results, as currently publicly reported, do not supply a key piece of information for our analysis— information about the order of the questions in the student's tests. Fortunately, in 2009 this information became freely available and we use this wave of data as the base of our analysis.

Although the main focus of PISA 2009 was reading comprehension, both math and science skills were also tested. The questions that measured all three of these areas were grouped around stimuli—texts, tables, charts, or some combination of these. In PISA 2009, 131 items were developed aimed to measure reading comprehension, which were grouped around 37 stimuli. Science comprehension items were grouped around 53 stimuli, and math comprehension around 34 stimuli. The average total time a student would need to complete all of these questions was estimated to be approximately seven-and-a half hours. Therefore, to shorten the time of student testing, it was decided that each student would be required to respond to only some of the questions on each of the areas tested.

To decide which question items a student should answer, first stimuli were grouped into thirteen blocks: seven blocks on reading, three blocks on science and three blocks on math. It was then estimated that the time required to answer each of these blocks would be around thirty minutes. After grouping stimuli in these blocks, thirteen different models of tests were developed, each containing four different blocks. Students were then presented with a randomly chosen version of the test from the thirteen versions that were prepared. This process resulted in tests that would require a maximum of two hours to be fully completed. In addition, as a result of the different versions of the questionnaire that were available, not all students responded to the same questions necessarily, or to questions in the same order, or even to the same proportion of questions about each of the areas of knowledge that were tested.

One essential aspect of the 2009 PISA test, which should be stressed, is the random assignment of tests to students, who were then randomly assigned one of the thirteen versions available. This unique feature of the provision of the PISA test will allow us to identify, after controlling for the necessary variables described in the next section, the effect of question order on the probability of answering correctly. This will be key for our measures of non-cognitive skills based on student's pattern of response to the PISA tests.

2.2 The Empirical Model

In our methodology, we first follow the work of Borghans and Schils (2012) and study how the probability of answering a question correctly decays as the test progresses. This effect could be understood as the rate of "test fatigue". For this aim, we use linear regression models of the following type:

$$y_{ij} = \delta_0 + \delta_1 Q_{ij} + \delta_2 X_i + \delta_3 CCAA_i + \delta_4 X_i * Q_{ij} + \delta_5 CCAA * Q_{ij} + \gamma_j + \varepsilon_{ij} \quad (1)$$

Where y_{ij} is 1 if the student responded correctly to question j, 0.5 if the student responded with a partially correct answer, and 0 if the answer was incorrect. Q_{ij} is the order of sequence of question j in the questionnaire that student i responded, re-scaled such that 0 becomes the first question and 1 becomes the last question. Its estimated coefficient δ_1 captures the rate of decline in performance as the test progresses or what we called, "test fatigue". X_i represents relevant personal variables of the student, the student's family, or the school the student attends, which we control for in our analysis. These variables include: age and gender of the student, father's and mother's highest level of education, father's and mother's occupation type, a variable indicative of whether the school is private, information about the size of the city where the school is located, number of books in the student's home, language(s) spoken in the student's home, shortage of teachers in the student's school, and the average economic and social level of students attending the school. Additionally, $CCAA_i$ includes variables indicative of the different autonomous communities of Spain that participated in PISA 2009.

We also included interaction terms between the order of sequence variables (Q_{ij}) and both the Spanish region dummies ($CCAA_i$) and the explanatory variables in X_i. This is toallow for the rate of "test fatigue" to be different for students with different characteristics or living in different regions in Spain. Finally, we include question fixed effects (γ_j), in order to control for the different levels of difficulty of questions. The error term ε_{ij} collects any unobservable factors that can affect the probability of answering a question correctly.

In a second analysis, we follow the specification presented in (1) but, following the ideas of Hitt, Trivitt and Cheng (2015), we studied the determinants of the probability of leaving a question blank, conditioning on the overall position of the question in the PISA test. This analysis is important as it allows us to study the possible sources of our estimated rate of decline in performance as the test progresses or what we called rate of "test fatigue". Thus for this model, the dependent variable is defined so it takes a value of 1 for those questions that the student left blank and zero otherwise.

Finally, we also estimated models at the student level, and not at the student-question level as above. This analysis allow us to study other aspects of student's non-cognitive skills that could also help explain possible sources of different rates of "test fatigue". In particular, we study the determinants of the probability that students were able to answer all the questions, rightly or wrongly, in the time initially established, and the probability that they managed to improve their relative position in the distribution of overall performance in the test after a rocky start (resilience). This second type of models take the following form:

$$y_i = \alpha_0 + \alpha_1 X_i + \alpha_2 CCAA_i + \gamma_B + \varepsilon_i \quad (2)$$

Where y_i is equal to 1 if the student has either failed to respond to all questions on the test on time, or if the student initially falls in the bottom of the distribution of results for the first set of test questions, but later falls into the top of the distribution for the test overall (the latter being our measure of resilience), and 0 otherwise. Given the nature of our binary dependent variable, equation (2) was estimated using binary choice probit models. The estimates of interest in this second type of models are the coefficients associated with the indicators for the different regions in Spain ($CCAA_i$). That is, with these models, we aim to assess if there are still regional differences in the student's likelihood of failing to complete the test or on the likelihood of showing resilience, even after controlling for student, family and school characteristics.

3. Results

The first column of Table 1 presents the results from the linear regression model presented in equation (1). The dependent variable is equal to 1 if the student correctly answered a given question, 0.5 if the student's answer was partially correct, and 0 otherwise. The results were similar when we coded the partially correct answers as either incorrect or correct.

<Insert Table 1 here>

As we mentioned previously, the explanatory variable "Position of the question" (Q_{ij}) captures the reduction in the probability of correctly answering —completely or partially—any given question, as a question moves one position forward in the questionnaire and it is an overall measure of "test fatigue". To the extent that we controlled for specific question fixed effects—capturing the intrinsic difficulty of each question—and that the order of questions was randomized in each questionnaire and the questionnaire type received by a student was also randomly assigned, using this approach we are able to identify the effect of presentation order of any given question on the likelihood of student's providing a correct response.

The coefficients associated with the dummies for each autonomous community of Spain describe the initial effect on the likelihood of students providing correct responses of attending school in a given region in Spain. That is, they capture the difference in education performance across regions in Spain. The coefficient associated with the interaction between the indicator of an autonomous region and the position of the question in the PISA test represent regional differences in Spain on the rate of decline in performance over the course of the test or the rate of "test fatigue". These two effects are estimated in the model after discounting the effect that student, family, and school characteristics may have on the probability of answering each question correctly. Also, the inclusion of individual questions' fixed effects ensures that inconsistent difficulty of the different questions does not affect our estimated effects of overall "test fatigue" or regional differences of "test fatigue".

Our analysis included data for those Spanish regions for which PISA 2009 provided a representative sample; that is, *all* except Valencia, Extremadura, and Castile-La Mancha. Also, it should be pointed out that Andalusia is the Spanish region of reference and, therefore, the estimated coefficients associated with other regional dummies indicate the deviation in the effect of interest from the estimate for Andalusia. The estimates for Andalusia are then captured by the constant term, for the overall regional effects, and by the variable indicating the position of the question, for the interactions between the region's fixed effects and the variable that reports the position of the question in the test.

The results we obtained confirmed that the region of residence of the students influences both their overall initial performance on the PISA exam, as well as the rate of decline in their performance as the test progresses. This is because all of the coefficients associated with indicators of Spanish regions, interacted or not with the variable indicating the position of the question on the test, are significantly different from zero at standard levels of statistical significance. Note, this result is obtained once we have discounted the effect of the characteristics of the students, their families, and their schools.

The results presented in the first column of Table 1 emphasize the importance of the student environment on the likelihood of answering a question correctly on the PISA test. Our results show that indeed, students' home environment is essential, both through the educational attainment and occupation of their parents, as well as through reading habits of the family (represented here by the number of books in the student's home). Moreover, the school environment also affects the probability of answering correctly. This is indicated by the positive and statistically significant coefficient associated with the variable that measures the average socioeconomic status of students surveyed on the PISA who attend the same school.

Additional estimates, not presented in Table 1 (to save space but available from the authors upon request), confirmed that the rate of "test fatigue" in PISA also is associated with certain student, family, and school characteristics, such as the language predominantly spoken in the student's home. The rate of "test fatigue" is reduced when the language is Spanish, the location of the school is in a larger city and when the average composition of the school is of higher socio-economic status. That is, we can say that the rate at which the probability of correctly answering a question in PISA declines is determined by context variables outside of the home, such as the autonomous region, the size of the town of residence, and the average socio-economic status of students in the school.

The estimates presented in the second column of Table 1 help us to further study the determinants of the differential rates of “test fatigue” for the PISA test. Specifically, these estimates analyze the relationship between the position of the question on the test and the likelihood that a student will leave a question blank. This possible relationship is relevant because of the evidence discussed in Hitt, Trivitt and Cheng (2015) who showed that the frequency with which students skip questions in a survey predicts ultimate educational attainment of these students, as well as their salaries later in life, even after controlling for their level of cognitive ability. Thus, we posit that the frequency with which students leave responses blank provides a measure of non-cognitive skills related to personality traits such as, perseverance or ability to concentrate. Hence, if we find that the decline in the rate of success over the course of the PISA test is correlated with the likelihood of leaving an answer blank, this will strengthen the interpretation previously presented in the literature, suggesting that these measures of “test fatigue” are measures of non-cognitive skills that lead to success in adult life.

The results in the second column of Table 1 suggest that the practice of leaving answers to a question blank is closely linked to the estimated rate of “test fatigue”, as it was presented in the first column of Table 1. For most regions, the estimated coefficient in the second column has the opposite sign to that estimated coefficient in the first column. This suggests that those regions with a higher tendency of skipping questions are ultimately those who present a higher rate of decline of correct answers as the PISA test progresses.

Our estimates suggest that students in those Spanish regions that after keeping student, family and school characteristics constant, have a greater overall likelihood of success in the PISA test are also less likely to leave a question blank. Also, those regions where the success rate declines faster over the course of the test (i.e. present higher rates of “test fatigue”) are also the most likely to have students leaving an answer blank.

The results presented thus far have analyzed the determinants of answering questions correctly on the PISA test, but do not inform on the relationship between the rate of “test fatigue” and the final score obtained by the student on the PISA test. Table 2 shows estimates using a linear regression model where the dependent variable is a student's final test score in the math PISA test, and the explanatory variables include student, family, and school characteristics, as was the case in the estimates presented in Table 1. In this model, however, we added two additional sets of variables constructed from the estimates presented in the first column of Table 1: the estimated overall effect of living in each autonomous community in Spain and the rate of “test fatigue” for each region. We also obtained results considering test scores in reading and science, but they were both qualitatively and quantitatively very similar to those obtained for math. Therefore, we decided to focus our discussion on math test scores for the sake of simplicity.

The results of the regression on student math scores confirm those previously found in the literature, where performance of Spanish students in the PISA test is predicted by student characteristics such as age and gender, as well as a wide range of household characteristics, such as parents' education levels, parents' occupational status, and the number of books in the home. The characteristics of the school are found to be much less relevant and statistically significant at conventional levels but only for the effect of the average socio-economic status of students in the school. We also find that the autonomous community of residence predicted student performance in math, just as much as the other relevant socio-economic factors described above, both through their overall regional differences on answering questions correctly, as well as through the different rates of “test fatigue”. Our results suggest that the first of these factors, namely regional differences in the initial success rate, is associated with regional differences in cognitive factors, not captured by the rest of explanatory variables included in the model. The second factor (rate of decline of performance during the course of the test or rate of “test fatigue”) captures relevant educational performance differences that are due to differences in certain non-cognitive skills or related personality traits that vary across Spanish regions. This latter result, suggests that the differences in student achievement observed among students from the different regions in Spain are due, in part, to regional differences in non-cognitive skills, such as perseverance, that have been found to lead to success in education and adult life outcomes (Heckman, 2011).

To get a better sense of the magnitude of previously presented estimated effects, we calculated the extent to which the variation in mathematics performance across the different Spanish autonomous communities would be reduced if regional differences in overall initial performance in the test or in the rates of “test fatigue” were reduced in a standard deviation. We found that a reduction of one standard deviation in regional differences on overall initial performance on the test would lead to a reduction in the dispersion of math test scores across Spanish regions of approximately 33%. Meanwhile, an equivalent reduction in regional differences in the rate of “test fatigue” would lead to a reduction in the regional dispersion of math test score performance of about 25%. That is, our estimates confirm that regional differences in non-cognitive skills account for an important share of the observed regional differences in PISA test scores. However, the share explained by non-cognitive skills is lower than that attributable to regional differences in cognitive dimensions or knowledge of the student. Together, these two factors explain almost 60% of the regional differences in average performance on the PISA test observed across the Spanish regions included in this study.

<Insert Table 2 here.>

In Table 3, we address the influence of the student's autonomous community of residence on two other important aspects of student’s pattern of response to the PISA test, potentially also related to non-cognitive abilities and not previously studied in the literature on the subject. First, we analyzed the probability that the student did not complete the test during the time allotment. Additionally, we studied the relationship between the student's region of residence and the likelihood that the student would improve his/her relative position in the distribution of student test scores following a rocky start in the test. The results obtained in this case were robust to different specifications changing the number of questions to account the beginning performance in the test as well as the percentiles used to measure the relative position of the student in the overall distribution of test performance.

<Insert Table 3 here.>

The results confirm the relationship between the region of residence and the likelihood that the student finishes covering all the questions in the PISA test questionnaire. We also find that the probability of a student improving his/her relative position in the distribution of all student test scores, after a rocky start, is clearly related to the student's autonomous community of residence. Thus, students who reside in regions, such as, La Rioja, Castile and Leon, Aragon, Galicia and Cantabria (presented in increasing order of magnitude of the estimated effect), and are otherwise similar on other individual, family, and school characteristics, are more likely to have had time to go over the entire PISA test questionnaire and present a higher probability of improving their relative position in the distribution of test scores for all students, after a difficult start. On the other hand, students who live in Andalusia, the Balearic Islands, and Ceuta and Melilla are less likely to improve their relative position in the distribution of all student test scores over the course of the PISA test, keeping other individual, family, and school characteristics constant. Note that these results are not incompatible with previous results presented above that showed a decline in test performance as the test progressed for most students.

Another interesting result that we found with these estimates is that there are very few family or school characteristics of the students that determine the probability of either being able to finish the questionnaire of the PISA test in the allotted time, or being able to improve their relative performance after a rocky start. This result suggests, once again, the importance of the territorial context as a determinant of performance in the PISA test. The only family variable that we found to have a significant effect was the number of books in a student's home. This variable was found to be positively related to both the probability of having finished answering the test as well as the probability of improving after a difficult start. Concerning school characteristics, only the average socio-economic level of the school was found to have a significant effect.

Summarizing so far, our results suggest a strong relationship between the regional rate of “test fatigue” in PISA 2009, the average proportion of students in the region that leave questions of the test blank, the average proportion that is able to answer correctly or not the whole questionnaire in the time provided, and the average proportion of students that are able to improve their relative position on the distribution of overall performance after suffering a rocky start in the test. These results suggest that these practices, which are closely related to student's perseverance, concentration, and intrinsic motivation, are an important part of the cause underlying the observed regional differences in performance in PISA across Spanish autonomous communities.

4. Discussion

The results presented in the tables above support the idea that a significant part of the observed regional differences in student performance in the PISA test in Spain are related to regional differences in the so-called non-cognitive skills. Our estimates suggest that measures of non-cognitive skills, related to certain personality traits such as perseverance or the ability to maintain effort and concentration, which are related to the ability to delay rewards and think long-term, differ significantly across students of different regions in Spain. These skills play an essential role in explaining regional differences in educational performance on the PISA test. These are the personality traits with potential to make a student more likely to maintain the effort over the course of the PISA test, or even improve his/her relative performance and ability to overcome an initially poor performance, which clearly demonstrates resilience.

Also, regional differences in non-cognitive skills also predicted the probability of students in different Spanish regions to be able to cover all questions of the PISA test in the allotted time and the probability of leaving questions blank. This latter behavior has been linked in the literature not only to lower performance on standardized tests, but also to: a lower likelihood of attaining higher levels of education, and lower working conditions and wages in the future. Hence, the estimated regional differences in non-cognitive skills among Spanish youth may explain, to a large extent, the observed regional differences in terms of educational attainment, dropout rates, working conditions and wellbeing in adulthood.

In particular, our estimates suggest that a reduction in a standard deviation of the variation across regions on the estimated rate of "test fatigue" would lead to a reduction of approximately 25% of the observed regional differences in performance in the PISA test. This appears to be a key variable for educational policy to the extent that an equivalent reduction in the regional differences of other relevant variables, such as the proportion of parents with the highest levels of education or the proportion of parents in top occupation categories, would lead to a much lower reduction, no more than 2%, of the observed differences in PISA scores across Spanish regions.

Our results are consistent with those obtained in other studies such as Cunha, Heckman, and Schnnach (2010) and Borghans, Meijers, and Weel (2008), which emphasize the importance of non-cognitive skills in explaining student performance. These studies found that non-cognitive skills and, in particular, those personality traits linked to perseverance, effort, and the ability to delay gratification are closely related to performance on standardized tests such as the PISA test. Our study corroborates the relevance of these non-cognitive skills in explaining student performance on the PISA test and in the resulting regional differences in student performance. Most importantly, we do so without having psychometric measures of non-cognitive skills, but by analyzing student response patterns on the PISA test. Finally, we go beyond the analysis of Borghans, Meijers, and Weel (2008) who studied the rate of decline on correct responses as the test progresses or what we have called rate of “test fatigue”. We further study other relevant patterns of response, such as the practice of leaving answers blank, not being able to complete the test in the allotted time frame, or being able to improve the performance in the course of the test after a difficult start.

Given the relevance of regional differences in non-cognitive skills in determining the observed differences in student performance in the PISA test, a logically important next question would be, whether is possible to promote the development of these non-cognitive skills in those regions where development of such skills is lagging behind. In light of the results in Mendez (2014), we think this might be possible through education policy interventions directed at teachers and parents to help promote these non-cognitive skills. While training for parents has, a priori, a higher expected return, teacher’s training may be easier to implement and standardize across the different regions in Spain and, therefore, this may be a more effective tool for education policy. More research is needed, however, to find the best design and assess the effectiveness of different possible interventions.

In this sense, the literature review conducted by the U.S Department of Education in 2013 provided a summary of those early educational interventions that were found to be a promising first steps in promoting important non-cognitive skills such as, tenacity and perseverance. In this report, interventions were grouped into five categories, which include: reading programs in schools to promote values; interventions that seek to change the mindset and strategies for student learning; alternative school models; informal learning programs and digital learning programs with useful tools for teachers. The common feature of these five types of programs is the promotion of new ways of learning and teaching. These are more individualized and focused on getting to know each particular student better in order to be able to enhance his/her strengths for lasting and permanent results. In this sense, Diamond and Lee (2011) reviewed empirical evidence of the effectiveness of training programs found to possibly help improve non-cognitive skills in primary and secondary school students, in particular through the promotion of executive function skills (i.e., those functions necessary in the development of the ability to concentrate, ignore distractions, retain and use new information, plan action and review plans established when necessary, and inhibit impulses). Among the programs that the authors featured in their article one can find computer-based training programs of non-cognitive skills, particular kinds of board games and physical exercise (aerobic, martial arts, yoga). However, overall, recent curriculum developments emphasize the development of specific non-cognitive abilities, such as perseverance, to the point of making them in some cases not only part of the curriculum but a separate subject with dedicated lecturing time and for which students are graded.

It should be pointed out that, to this date, none of these initiatives have been developed in Spain. Therefore, we believe that the next step for Spain, would be to learn, from the specifics of its own educational system and those proposed for other countries, to then decide which innovative interventions, would be successful approaches to adopt in Spain, based on experimental evaluations. Lessons from these experiments would help ensure the wide-spread success of future programs, reaching larger student populations, could ensure that Spanish students are able to gain the necessary skills for their educational success, optimal health and best possible performance in the labor market in adulthood. The extent to which context is essential to ensure the replication of the results of innovative educational interventions, it would be unreasonable to embrace some of the successful interventions in other countries like the United States without having carried out the necessary experimental analysis within the Spanish educational context to confirm the suitability of the same interventions.. Additionally, evidence from other countries is limited and more experimental research is needed to better understand which innovative interventions could help promote important non-cognitive skills.

Finally, it should be noted that the authors of this article are working with the School Board in the Region of Murcia to implement an innovative educational program that aims to improve non-cognitive skills in primary school students within that region of Spain. This innovative program, the first of its kind in Spain, aims in the short-term to improve student's achievement, while in the medium - and long-term aims to reduce the rate of students being held back, drop-out rates, social exclusion, crime, and teenage pregnancy, all variables that non-cognitive skills have been shown to influence.

5. Conclusions

The aim of this paper was to analyze the role of student's non-cognitive skills and, in particular, to determine observed regional differences of student performance on the 2009 PISA test, across the different autonomous communities of Spain. To date, with the exception of Balart and Cabrales (2014), the debate on the origin of these regional differences on student performance has been confined to the study of differences in family and school contexts for impacting student achievement. These previous studies, however, tended to find that a significant proportion of the observed regional differences remained unexplained after accounting for differences in families and in schools.

Using data from the PISA 2009 study, we found that a significant part of the observed regional differences in student's performance in Spain were due to regional differences in the level of non-cognitive skills of students across regions. In the absence of direct measurements of non-cognitive skills of students, we used performance tasks measures based on student's patterns of response to the questionnaire test, some of which were already validated in the literature. These include the rate of decline in test performance as the test progresses or what we called rate of "test fatigue", the tendency to leave questions blank, along with newly developed measures like the capacity to at least try to respond to all questions (as many as 64 questions) on the test in the time given, or being able to improve the relative position in the distribution of test scores after a rocky start in the test. We believe that these measures reflect non-cognitive skills related to individual traits such as concentration, perseverance, resilience, etc., as well as to other aspects of personality, which are positively associated with educational success, as it has been shown to be the case for those measures that have been validated. Our contribution to this literature is then, the finding of regional patterns of student differences in these skills and the quantifying of their effect on observed differences in student performance.

Specifically, our estimates suggest that a reduction of one standard deviation in the regional variation on the rate of "test fatigue" on the PISA test would lead to a reduction of approximately 25% of the regional variation on observed student performance in the PISA test. We believe this is a key variable for educational policy to the extent that an equivalent reduction in the regional differences of other relevant variables, such as the level of education or quality of occupations of students' parents, would be associated with a reduction no larger than 2% of the observed differences in average performance in PISA across the different Spanish regions.

Finally, we also find considerable differences across regions in Spain on the probability of students being able to organize themselves to cover the entire test questionnaire. Significant regional differences are also obtained when studying student's tendency to leave questions blank. This latter behavior has been linked in the recent literature, not only to lower performance on standardized PISA-type tests, but also to a decreased likelihood of achieving higher levels of education in adulthood and of having a good occupation and wages in the future. Therefore, we believe that the estimated regional differences in non-cognitive skills among young students may explain, to a large extent, the observed differences in educational attainment, dropout rates, working conditions and wellbeing in adulthood, across different regions in Spain.

6. References

BORGHANS, L., y SCHILS, T. (2012): The leaning tower of Pisa. Manuscipt. http://www.sole-jole.org/13260.pdf

BORGHANS, L., MEIJERS, H., WEEL, B.T. (2008): The role of noncognitive skills in explaining cognitive test scores. *Economic Inquiry*, vol. *46(1)*: 2-12.

COLEMAN, J. S., CAMPBELL, E. Q., HOBSON, C. J., MCPARTLAND, J., MOOD, A. M., WEINFELD, F.D., & YORK, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

CORDERO, J.M., CRESPO, E., PEDRAJA, F. (2013): Rendimiento educativo y determinantes según PISA: una revisión de la literatura en España. *Revista de Educación*, pp 273-297.

CUNHA, F., HECKMAN, J.J., SCHENNACH, S.M. (2010): Estimating the Technology of Cognitive and Noncognitive Skill Formation, Econometrica, Econometric Society, vol. 78(3): 883-931.

DEPARTAMENTO DE EDUCACIÓN DE LOS ESTADOS UNIDOS (2013): Promoting Grit, Tenacity, and Perseverance: Critical Factors for Success in the 21st Century. Office of Educational Technology.

DIAMOND, A., y LEE, K. (2011): Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old. *Science*, vol. *333*: 959-964.

GIL ESCUDERO, G. y SANZ LABRADOR, I. (2014): PISA para Centros Educativos: Un instrumento para el fomento del éxito escolar de los centros. *Participación Educativa*, pp 29-37

GONZÁLEZ SAN ROMÁN, A. DE LA RICA GOIRICELAYA, SARA (2012): Determinantes de las diferencias regionales en el rendimiento académico en España-PISA 2009. En *Educación y desarrollo: PISA 2009 y el sistema educativo español*. Fundación BBVA

HECKMAN, J.J. (2011): Integrating Personality Psychology into Economics. IZA Discussion Paper 5950, Institute for the Study of Labor (IZA).

HECKMAN, J. J., STIXRUD, J., y URZUA, S. (2006): The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24(3)*, 411-482.

HITT, C., TRIVITT, J., y CHENG, A. (2015): When you say nothing at all: The surprisingly predictive power of student effort on surveys. *EDRE Working paper*, http://www.uaedreform.org/downloads/2014/10/edre-wp-2014-10.pdf.

MENDEZ, I. (2014): Habilidades no cognitivas y rendimiento escolar. Avances en Supervisión Educativa 20.

VILLAR, A. (2012): PISA 2009 y el sistema educativo español: una visión panorámica. En *Educación y desarrollo: PISA 2009 y el sistema educativo español*. Fundación BBVA

Table 1: Determinants of Test fatigue Measures

Explanatory Variable	Variable to explain: probability of Answering Correctly	Leaving Blank
Position of the question	-0.171***	0.078***
	(0.001)	(0.000)
Fixed Effects of the Autonomous Communities		
Aragon	0.004*	-0.010***
	(0.002)	(0.001)
Asturias	0.006***	0.005***
	(0.002)	(0.001)
Balearic Islands	-0.028***	0.018***
	(0.004)	(0.002)
Canary Islands	-0.005***	0.006***
	(0.002)	(0.001)
Cantabria	-0.005***	0.003***
	(0.002)	(0.001)
Castile and Leon	0.005**	-0.006***
	(0.003)	(0.002)
Catalonia	0.018***	-0.012***
	(0.005)	(0.002)
Galicia	0.030***	0.006***
	(0.006)	(0.002)
La Rioja	0.015***	-0.003**
	(0.002)	(0.002)
Madrid	-0.005***	-0.004**
	(0.002)	(0.001)
Murcia	0.005***	-0.003***
	(0.002)	(0.001)
Navarra	0.018***	-0.014***
	(0.003)	(0.002)
Basque Country	-0.009***	-0.001
	(0.002)	(0.002)
Ceuta and Melilla	-0.076***	0.031***
	(0.003)	(0.001)

Table 1: Determinants of Test fatigue Measures (continuation)

Variable	Answering Correctly	Leaving Blank
Interactions with the question's position in the test		
Aragon	0.049***	-0.033***
	(0.000)	(0.000)
Asturias	0.035***	-0.010***
	(0.000)	(0.000)
Balearic Islands	-0.014***	0.011***
	(0.000)	(0.000)
Canary Islands	-0.036***	-0.017***
	(0.000)	(0.000)
Cantabria	0.052***	-0.030***
	(0.000)	(0.000)
Castile and Leon	0.062***	-0.035***
	(0.000)	(0.000)
Catalonia	0.008***	-0.021***
	(0.000)	(0.000)
Galicia	0.031***	-0.032***
	(0.000)	(0.000)
La Rioja	0.064***	-0.031***
	(0.000)	(0.000)
Madrid	0.058***	-0.035***
	(0.000)	(0.000)
Murcia	0.015***	-0.016***
	(0.000)	(0.000)
Navarra	0.028***	-0.019***
	(0.000)	(0.000)
Basque Country	0.040***	-0.024***
	(0.000)	(0.000)
Ceuta and Melilla	-0.002***	-0.012***
	(0.000)	(0.000)

Table 1: Determinants of Test fatigue Measures (continuation)

Variable	Answering Correctly	Leaving Blank
Student, Family, and School Characteristics		
Student age	0.021***	-0.005***
	(0.003)	(0.002)
Student gender (female)	0.003	-0.006**
	(0.005)	(0.002)
Highest level of education of the father	-0.005	0.003
	(0.005)	(0.004)
Average level of education of the father	-0.003	0.004**
	(0.005)	(0.002)
Highest level of education of the mother	0.001	-0.006**
	(0.007)	(0.003)
Average level of education of the mother	0.012**	-0.009***
	(0.006)	(0.003)
Father, occupation category 1	0.049***	-0.014*
	(0.013)	(0.007)
Father, occupation category 2	0.057***	-0.007***
	(0.005)	(0.002)
Father, occupation category 3	0.050***	-0.010***
	(0.003)	(0.001)
Father, occupation category 4	0.060***	-0.011***
	(0.006)	(0.004)
Father, occupation category 5	0.007	0.002
	(0.006)	(0.003)
Father, occupation category 6	-0.001	-0.011
	(0.017)	(0.008)
Father, occupation category 7	0.025*	-0.016***
	(0.013)	(0.004)
Father, occupation category 8	0.009	0.004
	(0.020)	(0.017)
Mother, occupation category 1	0.033***	-0.007
	(0.008)	(0.004)
Mother, occupation category 2	0.047***	-0.018***
	(0.006)	(0.004)
Mother, occupation category 3	0.048***	-0.018***
	(0.003)	(0.005)
Mother, occupation category 4	0.034***	-0.015***
	(0.005)	(0.002)
Mother, occupation category 5	0.016**	-0.009***
	(0.006)	(0.002)
Mother, occupation category 6	0.014*	-0.012*
	(0.008)	(0.007)
Mother, occupation category 7	0.007**	-0.007*
	(0.003)	(0.004)
Mother, occupation category 8	0.029***	-0.015***
	(0.009)	(0.002)

Table 1: Determinants of Test fatigue Measures (continuation)

Variable	Answering Correctly	Leaving Blank
Private School	-0.004	-0.002
	(0.010)	(0.004)
Medium City	0.004	0.000
	(0.003)	(0.003)
Large City	0.006	-0.001
	(0.008)	(0.006)
Number of books in the home		
11-25	0.062***	-0.032***
	(0.012)	(0.009)
26-100	0.118***	-0.053***
	(0.007)	(0.006)
101-200	0.164***	-0.063***
	(0.009)	(0.005)
201-500	0.186***	-0.068***
	(0.009)	(0.006)
More than 500	0.195***	-0.070***
	(0.012)	(0.007)
Language at home: Spanish	0.025**	-0.003
	(0.010)	(0.003)
Lack of teachers in School	0.003	-0.002
	(0.009)	(0.005)
Average school SES	0.040***	-0.022***
	(0.006)	(0.003)
Constant	0.063*	0.164***
	(0.038)	(0.032)
Sample size	1153820	1153820
Goodness of fit	0.1882	0.1091

Notes: Table 1 provides estimated OLS coefficients and robust standard errors adjusted for clustering at the level of autonomous community. The symbols *, ** and *** refer to the 10%, 5%, and 1% significance level of the estimated coefficients. The occupational categories are: business administration and public administration (1); technical and scientific professionals and intellectuals (2); technical and professional support (3); administrative employees (4); employees in food service, personal service, security, and sales (5); Skilled agricultural and fishery workers (6); craftsmen and skilled workers (7); plant and machine operators and assemblers (8). The reference category is unskilled workers. Medium- and large-sized cities have between 15,000 and 100,000 inhabitants and more than 100,000 inhabitants, respectively. The variable "lack of teachers" indicates if, in the school principal's view, the school exhibited a shortage of qualified math, science, or language school teachers.

Table 2. The effect of regional differences in cognitive and non-cognitive aspects on student math performance.

Variable	Score Mathematics
Fixed Effect of Autonomous Community	417.567***
	(112.522)
Interaction of Autonomous Community and Question Position	245.831***
	(74.625)
Characteristics of the student, family, and school	
Student age	6.814**
	(3.028)
Student gender (female)	-22.522***
	(2.635)
Highest level of education of the father	-1.673
	(1.987)
Average level of education of the father	0.660
	(1.824)
Highest level of education of the mother	7.339**
	(3.659)
Average level of education of the mother	12.155***
	(4.563)
Father, occupation category 1	21.699***
	(5.686)
Father, occupation category 2	26.396***
	(3.207)
Father, occupation category 3	20.555***
	(3.551)
Father, occupation category 4	23.820***
	(2.817)
Father, occupation category 5	3.794*
	(2.249)
Father, occupation category 6	6.561
	(8.345)
Father, occupation category 7	12.031***
	(4.327)
Father, occupation category 8	4.102
	(8.083)
Mother, occupation category 1	14.357***
	(3.910)
Mother, occupation category 2	25.205***
	(4.044)

Table 2. The effect of regional differences in cognitive and non-cognitive aspects on student math performance (continuation)

Variable	Score Mathematics
Mother, occupation category 3	18.647***
	(2.510)
Mother, occupation category 4	17.075***
	(2.833)
Mother, occupation category 5	6.002*
	(3.413)
Mother, occupation category 6	8.077**
	(3.164)
Mother, occupation category 7	3.283
	(2.272)
Mother, occupation category 8	12.604***
	(3.569)
Private School	-5.619
	(5.461)
Medium City	-0.460
	(1.292)
Large City	6.455**
	(3.292)
Number of books in the home	
11-25	37.381***
	(4.930)
26-100	66.630***
	(3.032)
101-200	92.382***
	(4.182)
201-500	100.177***
	(3.774)
More than 500	101.285***
	(4.374)
Language at home: Spanish	4.874**
	(2.189)
Lack of teachers in School	0.062
	(2.735)
Average school SES	16.500***
	(1.315)
Constant	296.159***
	(43.119)
Sample size	20284
Goodness of fit	0.307

Notes: Table 2 provides estimated OLS coefficients and robust standard errors adjusted for clustering at the level of autonomous community. The symbols *, ** and *** refer to the 10%, 5%, and 1% significance level of the estimated coefficient. Variables defined as in Table 1.

Table 3. Determinants of a student's probability of answering all questions on the test and improving his/her relative position in the distribution of test scores as the test progresses.

Fixed Effects of the Autonomous Communities	Did not complete exam	Resilience
Aragon	-0.386***	0.200***
	(0.009)	(0.071)
Asturias	-0.374***	0.190***
	(0.011)	(0.071)
Balearic Islands	0.115***	-0.184**
	(0.024)	(0.084)
Canary Islands	0.247***	-0.058
	(0.014)	(0.083)
Cantabria	-0.313***	0.199***
	(0.008)	(0.075)
Castile and Leon	-0.385***	0.299***
	(0.014)	(0.070)
Catalonia	-0.011	0.169**
	(0.026)	(0.079)
Galicia	-0.327***	0.219***
	(0.026)	(0.080)
La Rioja	-0.507***	0.293***
	(0.013)	(0.077)
Madrid	-0.373***	0.111
	(0.013)	(0.073)
Murcia	-0.117***	0.110
	(0.008)	(0.077)
Navarra	-0.163***	0.181**
	(0.018)	(0.089)
Basque Country	-0.353***	0.098*
	(0.010)	(0.058)
Ceuta and Melilla	0.152***	-0.143
	(0.025)	(0.091)
Student, Family, and School Characteristics		
Student age	0.153***	0.009
	(0.042)	(0.076)
Student gender (female)	0.166***	0.016
	(0.016)	(0.043)

Table 3. Determinants of a student's probability of answering all questions on the test and improving his/her relative position in the distribution of test scores as the test progresses (continuation)

Student, Family, and School Characteristics	Did not complete exam	Resilience
Highest level of education of the father	0.076***	0.069
	(0.021)	(0.083)
Average level of education of the father	0.020	0.030
	(0.031)	(0.073)
Highest level of education of the mother	-0.045	-0.137
	(0.054)	(0.090)
Average level of education of the mother	-0.070***	-0.053
	(0.025)	(0.077)
Father, occupation category 1	0.081	-0.023
	(0.087)	(0.115)
Father, occupation category 2	-0.022	0.118
	(0.043)	(0.084)
Father, occupation category 3	0.054	0.142*
	(0.059)	(0.079)
Father, occupation category 4	0.014	0.113
	(0.072)	(0.073)
Father, occupation category 5	0.138***	-0.079
	(0.051)	(0.064)
Father, occupation category 6	0.351	-0.325
	(0.221)	(0.201)
Father, occupation category 7	0.009	-0.056
	(0.089)	(0.108)
Father, occupation category 8	0.036	-0.157
	(0.093)	(0.159)
Mother, occupation category 1	-0.099**	-0.015
	(0.040)	(0.091)
Mother, occupation category 2	-0.042	0.071
	(0.070)	(0.099)
Mother, occupation category 3	0.024	0.109
	(0.074)	(0.098)
Mother, occupation category 4	0.022	0.100
	(0.073)	(0.107)
Mother, occupation category 5	-0.092*	0.042
	(0.053)	(0.094)
Mother, occupation category 6	0.071	0.017
	(0.067)	(0.152)
Mother, occupation category 7	-0.025	-0.137*
	(0.076)	(0.079)
Mother, occupation category 8	0.036	0.051
	(0.061)	(0.095)

Table 3. Determinants of a student's probability of answering all questions on the test and improving his/her relative position in the distribution of test scores as the test progresses (continuation)

Student, Family, and School Characteristics	Did not complete exam	Resilience
Private School	0.018	-0.004
	(0.042)	(0.061)
Medium City	-0.022	-0.037
	(0.060)	(0.055)
Large City	-0.009	-0.007
	(0.023)	(0.053)
Number of books in the home		
11-25	-0.177**	-0.173
	(0.069)	(0.132)
26-100	-0.232***	0.169
	(0.044)	(0.122)
101-200	-0.262***	0.312**
	(0.049)	(0.125)
201-500	-0.328***	0.457***
	(0.062)	(0.129)
More than 500	-0.392***	0.587***
	(0.089)	(0.135)
Language at home: Spanish	0.034	0.150**
	(0.040)	(0.072)
Lack of teachers in School	0.068	-0.056
	(0.083)	(0.049)
Average school SES	-0.131***	0.168***
	(0.028)	(0.044)
Constant	-3.249***	-1.717
	(0.628)	(1.223)
Sample size	20284	12806
Goodness of fit	0.0412	0.1699

Notes: Table 3 provides estimated coefficients using a discrete choice probit model with robust standard errors adjusted for clustering at the level of autonomous community. The symbols *, ** and *** refer to the 10%, 5%, and 1% significance level of the estimated coefficient. Variables defined as in Table 1.