CATÓLICA
CATÓLICA PORTO BUSINESS SCHOOL
PORTO

WORKING PAPERS

MANAGEMENT

Nº 01/2016

STUDENT PERFORMANCE IN MATHEMATICS USING PISA-2009 DATA FOR PORTUGAL

Susana Faria
Universidade do Minho

M. Conceição Portela
Católica Porto Business School

Student Performance in Mathematics using PISA-2009 data for Portugal

September 28, 2016

Susana Faria[a] and Maria C.S. Portela[b]

[a]*CMAT - Centre of Mathematics, Department of Mathematics and Applications, University of Minho, Portugal*;
[b]*CEGE - Centro de Estudos em Gestão e Economia da Católica Porto Business School, Portugal*

Abstract

This paper is based on Portuguese data from PISA-2009, and it focuses on the measurement of student achievement in mathematics and on the determinants of this achievement both at the student and at the school levels. Data on about 3900 Portuguese students and 194 schools who participated in PISA-2009 were used to accomplish our objectives. Given the hierarchical structure of data, the models adopted for statistical analysis were multilevel models, which can take into account data variability within and among the hierarchical levels. Specifically we were interested in understanding whether the impact of students' variables were similar for students with different levels of achievement. As a result, we used a multilevel quantile regression model to analyse the determinants of students' success, where the potential determinants are student and school variables. Our study provides evidence that a stable relation with achievement is expected for some variables (e.g. gender, repetition, or socio economic background), while other variables show varying impacts depending on the students location on the rank of achievement in maths (e.g. immigrant status of students, or some study strategies like control strategies). In spite of schools having a significant impact on students' achievement (without considering any explanatory factors, 30% of the variability found in students' test scores can be explained by the school attended), we found that most school-level variables (except location) were not significant in explaining the school effect.

Keywords: Multilevel quantile regression models; Mathematics achievements; PISA-2009

1 Introduction

The development of international tests like PISA, TIMSS or PIRLS has been a major achievement of education scientists and institutions. Cross-country analysis allows one to take account of institutional and political effects on students achievements, that could not be tested with country specific data. In an analysis of TIMSS data on 39 countries and more than 260,000 students, Woessmann Woessmann (2003) estimated, among other things, the percentage of variation in student results that is due to the class, the school and the country, both for mathematics and science. Results suggested that the highest percentage of total variation is due to classes (30% for science and 48.6% for mathematics), followed by schools (40.5% in mathematics and 29.5% is science) and finally by the country (25.1% in mathematics and 16.2% in science). Schools and countries play therefore a considerable role in explaining differences in achievement of students, but classes are usually more heterogeneous than schools themselves. Several authors, like Woessmann Woessmann (2003), Hanchane and Mostafa Hanchane and Mostafa (2012), or Hanushek et al. Hanushek et al. (2013) have used international databases to analyse cross-country differences and the impact on students achievements of educational policies (e.g. the existence of central examinations, different levels of stratification, the autonomy of schools, their ownership, etc.).

Other authors have used international data sets to perform country-specific analysis, since in many countries national exam datasets have a limited access or are not as rich as international datasets. This is the case of Portugal, where there are publicly available data on the results of students in national exams, but data on the socio-economic background of students are virtually non-existent for academic research. Portugal is, as a result a country where the use of international data sets, such as PISA may be the only means to access the impact of several variables (regarding pupils, schools or even the country educational policies) on the achievement of Portuguese students. As the recent OECD (Organization for Economic Co-operation and Development) reveals, Portugal is amongst the lowest ranked countries in terms of percentage of the population aged between 25 and 34 and between 55 and 64 that has achieved upper secondary education. In addition, "more than 40% of young people from low educational backgrounds have not completed upper secondary education, and less than 20% of those young people have enrolled in tertiary education" (OECD, 2012, p.108).

There are some conceptual models in the literature to explain students' achievement or educational effectiveness. The one that served as a basis for this study is the 'educational production function' model as described in Hanushek (1979). This conceptual model establishes that students' achievement depends on inputs grouped into four main dimensions: Family background, Peer influences, School inputs, and Innate abilities. These inputs are organized in different levels, and as a result multilevel statistical models have been the main instrument of analysis in 'educational production functions' approaches (examples can be found in Agasisti et al. (2014); Ladd and Walsh (2002); Hanushek and Taylor (1990); Goldstein et al. (1993); Gray and Jesson (1996); O'Donoghue et al. (1997)). Typically multilevel models allow the derivation of the variables that more strongly determine the achievement of students and also the derivation of school effects. These school effects can be interpreted as value-added measures when the variables considered in the analysis take into account previous attainment, which is used to explain current attainment (OECD, 2008). Note that depending on the variables included in the model, school effects may have different interpretations. As pointed out by Raudenbush and Willms (Raudenbush and Willms, 1995, p.308) school effects may be interpreted as "the extent to which attending a particular school modifies a student outcome", or as "the effect on a student outcome of a particular policy or practice, such as the effect of reducing a student-teacher ratio or the effect of adopting a school-wide peer tutoring program". The former perspective corresponds usually to the measurement of value-added, while the latter perspective corresponds to the assessment of school efficiency. This distinction between efficiency and value-added is also mentioned in Hanushek (1979).

This paper adopts a value added (VA) perspective of school effects, and a multilevel model is used to analyse the impact of student and school factors on mathematics achievement of Portuguese students in PISA 2009. We depart, however, from extant literature by considering a multilevel quantile regression model. This is justified, by our concern of understanding whereas the impact of certain variables on students test scores is similar across different levels of student achievements. In addition, school effects may also be different across different levels of achievement, as deliberately some schools devise strategies more focused on low achieving students (like the creation of special classes, additional teaching time, etc.) while others are more focused on high achieving students (those that tend to focus more on average results than on success rates, and tend to have additional classes for preparing students for exams). Therefore, in this paper we address the research question of finding how heterogeneous are the effects of schools and students' characteristics on determining Portuguese student mathematic scores in PISA 2009.

The focus on mathematics is justified by the fact that the determinants of achievement are not necessarily the same for all subjects. For example, Fuchs and Woessmann Fuchs and Woessmann (2004) report results on 174 227 students from 31 countries and show different coefficients estimated for reading, science and mathematics Pisa 2000 scores. In Woessmann (2003) the author reports consistent results for maths and science but some effects were more prevalent in one subject than the other. As a result, we decided to focus on a single subject, meaning that conclusions are not generalized to situations where other subjects are being used, since in that case the relationship between explanatory variables and test scores may be different. Our interest for mathematics is justified by the typical low achievement of Portuguese students in this subject (this is shown in PISA results where on average Portuguese students score less in mathematics than in the other subjects and also nationally where in the 9th year national exams students score on average less in maths than in reading, and later in the secondary examinations typically mathematics and physics are the only subjects with an average score below 10 (on a 0-20 scale)). The focus on PISA 2009 data rather than 2012 is related with the set of variables that we wished to include in the assessment that are not available (or had too many missing values) for the most recent pisa dataset (e.g. variables relating to approaches to learning).

There are several applications in the literature of multilevel models to PISA datasets, but not many analysing Portuguese students and even less applying multilevel quantile models as we do in this paper. Examples of country specific analysis using PISA results can be found in Agasisti and Cordero-Ferrera (2013), where Italian and Spanish students were analysed in PISA 2006 through multilevel models, in Alacaci and Erbas (2010), where Turkish students were analysed, also through multilevel models and using the same PISA dataset, or in Mancebón et al. (2012) who analysed Spanish students's science scores in PISA2006 through multilevel models. The analysis of PISA datasets at country level reveal very different determinants of students achievement and different importance of schools and school-specific variables. For example, in Turkey the between school variation accounts for 55% of total variation in students results, whereas in Italy and Spain these values are around 40% and 13%, respectively. Regional differences also account for a large proportion of variation in Italy (see also Agasisti et al. (2014) on this topic), while in Spain regional differences account for a small proportion of variance. Analysing PISA-2003 mathematics test scores, Martins and Veiga Martins and Veiga (2010) show that in Portugal 37% of the variation in students' achievement is explained by school differences, whereas a value around 60% is found for Austria, Belgium Netherlands and Germany, while for Finland it is only 5%, showing a high homogeneity of schools in this country. PISA data have also been used for country specific analysis using other methodologies. For example, Perelman and Santin Perelman and Santin (2011) analysed PISA2003 maths and reading scores for Spanish pupils though stochastic frontier methods, and Aristovnik and Obadic Aristovnik and Obadic (2014) applied frontier models (non-parametric ones) to PISA2006 results but analysed school level data rather than pupil level data. Kruger Kruger (2011) also analysed school aggregated data from PISA 2009 addressing the degree of segmentation prevailing in the Argentine school system. Giambona and Porcu Giambona and Porcu (2015) analysed Italian students' performance in PISA2009 reading using a quantile regression, but they did not use a multilevel model.

Previous applications of multilevel linear regression models to the Portuguese context can be seen in Ferrão and Goldstein (2009) and Ferrão (2009). These studies, however, did not use PISA data, but student achievement on a specific mathematics test. To the authors' knowledge the only published studies analysing PISA data on Portuguese students through multilevel models are those of Dias and Ferrão (2006), Carneiro (2008), Valente et al. (2011), and Fonseca et al. (2011). From these studies we can infer that Portuguese schools explain around 35% of the variation found in students' scores in PISA. A figure of 34% was found in Dias and Ferrão (2006) when analysing PISA 2000 mathematics scores, and a figure of 37% was found by Martins and Veiga (2010) when analysing PISA 2003 mathematics scores. Some other previous studies focused on science achievements, like Fonseca et al. (2011) who used PISA 2006 results and analysed their relationship with attitudes of students towards science and Socio-economic and cultural status (ESCS), or Valente et al. (2011), who used the same data set to analyse the impact of teaching and learning strategies on students' science achievement. Other examples of interest applied to the Portuguese context can be found in Pereira and Reis (2014) who analysed student's retention in Portugal through a treatment effects model, based on PISA 2003 and 2009 data, or in Pereira (2011) who analysed the evolution of the performance of Portuguese students in PISA 2003, 2006 and 2009 tests (these authors have used standard regression models and quantile regression models but did not take into account the hierarchical structure of the data).

In summary, there are not many studies in the Portuguese context, applying multilevel models, and none applying multilevel quantile models as reported in this paper. In the international context there are also not many examples of studies applying multilevel quantile regression models. Some exceptions can be found in Geraci and Bottai (2014) who used an illustrative example on education, or Tian and Chen (2006) who report an application of a multilevel quantile model of mathematics achievements of Canadian students, or Costanzo (2015) who used a multilevel quantile regression model to analyse mathematics achievements of Italian students.

2 Methodology

Quantile regression (QR) estimates the conditional quantiles of a response variable distribution through a linear regression model and provides a more complete view of the relationships between variables. Since it was introduced by Koenker and Bassett Koenker and Bassett (1878) as a robust (to outliers) and flexible (about error distribution) linear regression method, it has received considerable interest in both theoretical and applied statistics. For an overview of recent applications of quantile regression see Komunjer (2005).

Consider data in the form $(\mathbf{x}_i^T, y_i), i = 1, \ldots, N$, where $\mathbf{x_i}^T$ are row $p-$vectors of a known design matrix $\mathbf{X}$ and y_i is a scalar response variable with conditional cumulative distribution function $F_{Y_i|x_i}$. In quantile regression problems, the objective is to estimate models of the type

$$Q_{Y_i|x_i}(\tau) = \mathbf{x}_i^T\beta^\tau, \quad i = 1, \ldots, N \tag{1}$$

where $Q_{Y_i|x_i} \equiv F^{-1}_{Y_i|x_i}$ denotes the inverse of the distribution function $F_{Y_i|x_i}$, $\tau \in (0, 1)$ denotes the quantile level of interest, and $\beta^\tau \in R^p$ is a column vector of unknown fixed regression coefficients. Alternatively, the equation (1) can be written as

$$y_i = \mathbf{x_i^T}\beta^\tau + \epsilon_i, \quad \text{with} \quad Q_{\epsilon_i|x_i}(\tau) = 0. \tag{2}$$

Given a random sample $(\mathbf{x}_i^T, y_i), i = 1, \ldots, N$, the estimator $\hat{\beta}^\tau$ is obtained by solving

$$\hat{\beta}^\tau = \underset{\beta\in R}{\operatorname{argmin}} \sum_{i=1}^{M} \rho_\tau(y_i - \mathbf{x_i^T}\beta^\tau) \tag{3}$$

where $\rho_\tau(\nu) = \nu\{\tau - I(\nu < 0)\}$ is the loss function, ν is a real number, and $I(.)$ is an indicator function. The loss function is a piecewise linear function that assigns weights τ to positive residuals and $(1 - \tau)$ to negative residuals.

A number of QR methods are based on the Asymmetric Laplace (AL) distribution. A continuous random variable $w \in R$ is said to follow an AL density with parameters (μ, σ, τ), $w \sim AL(\mu, \sigma, \tau)$, if its probability density can be expressed as

$$f(w|\mu,\sigma,\tau) = \frac{\tau(1-\tau)}{\sigma} exp\{-\frac{1}{\sigma}\rho_\tau(w-\mu)\} \tag{4}$$

where $-\infty < \mu < +\infty$ is the location parameter, $\sigma > 0$ is the scale parameter, and τ is the skewness parameter. See Yu and Zhang (2005) for more details.

Recently, Geraci and Bottai Geraci and Bottai (2014) have introduced a new method for quantile regression with mixed effects called linear quantile mixed models. They propose a conditional quantile regression model for continuous responses where random effects are added to the model taking into account the dependence between units when an hierarchical data structure is present.

We have adopted the procedure proposed by Geraci and Bottai (2014) to perform a multilevel quantile regression model.

Assume that we have data from J schools, each with a different number of students n_j. Consider data in the form $(\mathbf{x}_{ij}^T, y_{ij})$ for $i = 1,\ldots,n_j$ and $j = 1,\ldots,J$, $N = \sum_{j=1}^{J} n_j$, where $\mathbf{x}_{ij}^T$ is a vector of the of student level variables (student i attending school j) and y_{ij} are student PISA scores in mathematics. We considered the random intercept quantile regression model

$$y_{ij} = \beta_{0j}^\tau + \mathbf{x}_{ij}^T\beta^\tau + \varepsilon_{ij} \tag{5}$$

$$\beta_{0j}^\tau = \gamma_{00}^\tau + u_{0j} \tag{6}$$

where β^τ is a vector of unknown fixed effects, β_{0j}^τ is the intercept representing the average achievement for the j school (this intercept varies at the school level), γ_{00}^τ is the average achievement of the school means and $u_{0j}, j = 1,\ldots,J$ is the random effect associated with school j. The dependence among the students within the $j-$th school is induced by the random effect u_{0j} which is shared by all students within the same school. We assume that y_{ij} conditionally on u_{0j} are independently distributed according to an AL distribution with location and scale parameters given by $\mu_{ij} = \mathbf{x}_{ij}^T\beta^\tau + u_{0j}$ and σ^τ. The skew parameter τ is set *a priori* and defines the quantile level to be estimated. Also, we assume that the random effects $u_{0j}, j = 1,\ldots,J$ are mutually independent and identically distributed according to some density $f(u_{0j}|\psi^{\tau_u})$, where ψ_u^τ is a scale parameter. We assume the ε_{ij} are independent and u_{0j} and ε_{ij} are independent of one another.

School level variables (W_j) can be introduced in the model for explaining school effects. In that sense the equation above becomes:

$$\beta_{0j}^\tau = \gamma_{00}^\tau + \mathbf{W}_j^T\gamma^\tau + u_{0j}. \tag{7}$$

In applying the multilevel quantile model first estimate school effects without taking into account school characteristics, and only after analysing school effects and differences between schools, we estimated the multilevel quantile model with school level variables. This second step estimation attempted to explain school effects and decide which school level variables reveal significant in explaining such effects.

We can define the intraclass correlation (ICC),i.e., the ratio of the variance of the random effects to the total variance,

$$ICC = \frac{(\psi_u^\tau)^2}{(\psi^\tau)^2 + (\psi_u^\tau)^2} \tag{8}$$

where $(\psi_u^\tau)^2$ is the variance of the random effects and $(\psi^\tau)^2$ is the variance of model's error term. To calculate the variance of model's error term see Geraci (2014).

The first attempt to fit quantile regression models with random intercepts led to the Monte Carlo EM algorithm (Geraci and Bottai, 2007), which can be computationally intensive and inefficient. A new approach based on a combination of gaussian quadrature approximations and non-smooth optimization algorithms has been proposed by Geraci and Bottai (2014) and implemented in the R package **lqmm** (Geraci (2014)). This was the package used to estimate our empirical model. In addition, we used the Akaike information criterion (AIC) to calculate the fit of each model. We also estimated a two level standard multilevel model using the package **nmle** (Pinheiro et al. (2014)) for the statistical programming environment R.

3 Data set

The Programme for International Student Assessment (PISA) is an international study, launched by the OECD in 1997, that assesses every three years 15-year-olds' skills in three key subjects (reading, mathematics and science). Math achievement of students is the outcome of interest in this paper and, following Hanushek (1979) this achievement is a function of Family Background, Innate abilities, Peer effects, and School inputs. The actual variables included in each of these categories are shown in Table 1. In addition to these variables gender was also included. School variables were classified into 3 groups: context, resources, and policies/strategies. The first group of school variables regards contingencies that schools cannot control, whereas resources and policies are variables on which schools can act upon to improve their student results. (In Table 1 we use small caps for student variables' codes and big caps for school variables' codes).

Table 1: Variables description

	Student Variables
Family background	Socio-economic and cultural index (escs) Family structure (fs): 1- Traditional; 2- Only one parent; 3- Other Immigrant status (imi): 1- No; 2-Yes Math tutoring outside school(tut):1- No; 2 -Yes Help in homework (hmw):1- No; 2- Yes Expectations on leaving the school (exp): 1- 9 grade; 2- 12 grade professional; 3- 12 grade general; 4-Undergraduate or postgraduate course Home possessions (hpos)
Innate Abilities	Repeating student (rep): 1-No; 2-Yes Grade of student (grade): 1- Third cicle; 2 - Secondary education (general); 3 - Secondary education (professional) Use of control strategies (cst) Use of elaboration strategies (est) Use of summary strategies (sst) Use of comprehension and remembering (cr) Use of memorizing strategies (mem)
	School Variables
Context	Average of socio-economic and cultural index (ESCSAVG) Type of school (TYPE): 1- Public; 2- Private with public funding; 3- Private Independent Location (LOC): 1- City;2 - Village, Dimension of the school (DIM) Percentage of girls (PG) Percentage of students with math's tutoring outside the school (PTUT)
Resources	Student teacher ratio (STR) Computers per student (C/S) Proportion of computers linked to the web (COMPWEB)
Politicies /Strategies	Student behaviour (STBEA) Teacher behaviour (TEABEA) Index of extra-curricular activities (EXTC)

3.1 Family Background

The literature is unanimous regarding the importance of including socio-economic characteristics of students into the analysis. After analysing some raw variables reflecting this construct, it was decided that the index 'escs' (student socio-economic and cultural status) constructed within PISA (through factor analysis) was indeed the variable that better reflected our construct. This variable was derived in PISA from three indices: (1) the highest occupational status of parents; (2) the highest educational level of parents in years of education according to ISCED, and (3) home possessions. Home possessions was also considered in the set of variables as a robustness check. In addition to 'escs' we also used the family structure of the student (which we recoded in three levels, where the traditional structure is the one where the student lives with mother and father), the immigrant or non-immigrant status of the student and family, the existence of private tutoring in mathematics outside the school, and the help parents provide in students' homework.

The inclusion of the variable private tutoring is related to our aim of estimating the impact on math test scores that the attendance of a particular school may have on students' achievements. This impact may not be reasonably estimated when students have private tutoring to enhance their grades (in the sense that the work that should have been done by schools, is being done by other entities outside the school). Therefore in order to better estimate the school effects, it is our belief that this variable should be considered (note than in other countries the effect of this variable may not be significant, but in Portugal it is known that above 30% of students in the 11, and 12 years of secondary education have private tutoring to enhance their grades on national exams) (Romão, 2012). There is evidence from other countries regarding the impact of private tutoring on exam achievement. One such recent example is the study by Berberoglu and Tansel (2014) regarding evidence from Turkey.

Regarding the variable 'help in homework', we also believe that this is an important factor behind the student family context, as it could reveal the family support towards the student work at home. This variable was collected from the Parents questionnaire. Within the family background of the student we also included students' expectations regarding their academic future achievements as these may impact their success (as testified for example by Suárez-Álvarez et al. (2014), when analysing Spanish students). Students' expectations on leaving school are related to the family context of the student as students whose parents achieved a high professional and economic status, will aim, in principle, to reach the level of their parents, and their family also expect them to do so.

3.2 Innate Abilities

Regarding Innate abilities of students, there is no direct measure of these in PISA database as no measure of prior attainment is available. As a result we tried to look for variables that could reflect the prior attainment of students. In particular we have information regarding students having or not having repeated one or more years at school, and also information regarding the grade attended[1]. In spite of many differences in innate abilities that may distinguish two students in the same grade and with the same number of previous repetitions, students attending different grades shall have different acquired skills, and students that repeated some year have a background of

[1] In Portugal education is divided into cycles. The first cycle comprises the first 4 years of education. The second cycle comprises the 5^{th} and 6^{th} years of education, the third cycle relates to the 7^{th}, 8^{th} and 9^{th} years of education and secondary education comprises the 10^{th}, 11^{th} and 12^{th} years. Education is compulsory until the 12^{th} year since 2009 (before that only the 9^{th} year was compulsory).

failure that may distinguish their attainment from those students that never failed. Different innate abilities, may also translate in different approaches towards learning. According to OECD (2010a) in the PISA 2009 framework the approaches to learning entail several strategies, including memorization strategies, control strategies, elaboration strategies, understanding and remembering, and summarizing strategies. Variables reflecting these strategies were used in our study as a way to enlarge the set of variables accounting for innate abilities of students (known to be the main predictor of achievement). In particular, the factor relating to 'Elaboration strategies' captures the way students relate new information and prior knowledge, use information learned at school outside the school, and relate learned materials and personal experiences. 'Control strategies' capture students' ability to figure out what they need to learn when they study, the ability to check if they understand what they read, their ability to figure out the concepts that they still did not understand, their ability to make sure that they remember the most important issues, and whether they look for additional information to clarify concepts. 'Memorization strategies' capture whether students try to memorize everything they read, whether they try to memorize details, and whether they do this by repeated reading. As reported in OECD (2010a), students that rely heavily on memorizing strategies tend to process little information apart from that memorized. Regarding 'comprehension and remembering strategies' they relate to the discussion of contents after reading, the use of students own words in summaries, whether students read quickly through a text, etc. 'Summarizing strategies' relate to the above in the sense that they try to assess the extent to which students are effective in summarizing information that they read (where less effective strategies would be copying as many sentences as possible from the text to summarize it, and most effective strategies would be to write down in their own words the most important parts of the text).

3.3 Peer effects

Peer effects have been captured by the average socio economic background status of the pupils attending the same school (see e.g. Perelman and Santin, 2011). This peer effect variable is observed at the school level and not at the pupil level, as students attending the same school with be subject to the same peer influences. As only one variable is used to reflect peer effects and it is a school level variable, we show it within school variables in Table 1.

3.4 School inputs

At the school level, most of the variables in Table 1 deserve no further explanations, except for those relating to policies or strategies followed by the school. At this level we included two variables relating to the teachers and students behaviour at the school (these variables are obtainable from the school questionnaires, and therefore students' and teachers' behaviour is that assessed by the school). The student behaviour is an aggregate of questions relating to "the extent to which learning is hindered by behaviours such as student absenteeism, the use of alcohol or illegal drugs, bullying, disruption of classes by students, and students lack of respect for teachers" (OECD, 2010b). The teacher behaviour is also an aggregate of questions relating to the extent to which school principals "perceived learning in their schools to be hindered by such factors as teachers' low expectations of students, poor student-teacher relations, absenteeism among teachers, staff resistance to change, teachers not meeting individual students' needs, teachers being too strict with students and students not being encouraged to achieve their full potential" (OECD, 2010b). For both variables higher values mean behaviours that are less disruptive. A variable capturing the existence of many extra-curricular activities at the school was also considered, as it can be an indicator of the environment lived at the school. The inclusion of this variable allows the investigation of whether more dynamic schools on creating opportunities for extra-curricular activities tend to perform better. Several other policy related school variables, could have been considered. However, the school data set had many missing values and we decided to include only those variables that existed for the majority of schools in our sample.

3.5 Descriptive analysis of the variables

Our sample contains 3900 pupils, with ages between 15.25 and 16.25 years old attending a total of 194 schools. The dependent variable used in our models is each of the 5 Plausible Values (PV) produced within the PISA database. We report some descriptive statistics for these PVs in Table 2, but in the remaining of the paper we will report average results obtained from the use of each of the PVs. From all plausible values in Table 2, we can say that average achievement in maths is about 505 points for the sampled students. Students in quartile 1 have more than 100 points less than students in quartile 3, indicating a high difference between the two quartiles.

Table 3 shows some descriptives of the student-level continuous variables considered in our model. Given our interest in analysing and distinguishing the achievement of students in various quantiles, we preset our descriptive

Table 2: Descriptive statistics for plausible values in maths

	Mean	St Dev	Q_1	Median	Q_3
PV1	504.75	84.42	448.52	504.60	566.22
PV2	505.04	85.09	446.18	505.77	565.36
PV3	504.21	84.90	446.96	504.21	563.72
PV4	505.11	84.37	446.49	505.38	564.58
PV5	503.93	84.52	444.63	504.60	563.72
PVMEAN	504.61	84.66	446.56	504.91	564.72

analysis for different groups of students per quantile of achievement in Math test scores (in particular 1^{st}, 2^{nd}, 3^{rd}, and 4^{th} quartiles).

Table 3: Descriptive statistics of student-level continuous variables per quartile of maths achievement

	Students in quartile 1			Students in quartile 2		
	Mean	St Dev	Median	Mean	St Dev	Median
escs	-0.92	0.92	-1.01	-0.49	1.07	-0.63
hpos	0.06	0.85	0.02	0.38	0.85	0.39
cst	-0.22	0.92	-0.29	0.07	0.95	-0.04
est	0.18	0.80	0.04	0.31	0.81	0.33
sst	-0.46	1.07	-0.33	0.07	1.00	0.51
cr	-0.56	1.00	-0.78	-0.08	1.01	-0.05
mem	-0.18	0.92	-0.21	-0.17	0.91	-0.21
	Students in quartile 3			Students in quartile 4		
	Mean	St Dev	Median	Mean	St Dev	Median
escs	-0.08	1.12	-0.17	0.50	1.18	0.50
hpos	0.61	0.88	0.59	0.95	0.89	1.01
cst	0.32	0.92	0.22	0.63	0.85	0.49
est	0.47	0.81	0.33	0.73	0.82	0.60
sst	0.33	0.83	0.51	0.60	0.69	0.51
cr	0.27	0.94	0.32	0.55	0.82	0.68
mem	-0.29	0.95	-0.21	-0.50	1.00	-0.54

It is interesting to note that students in different quartiles of maths achievement have very different characteristics. In particular students in the fourth quartile have higher medians of escs (i.e. more favourable family context and socio-economic and cultural index) and higher medians of hpos, they use more intensively control strategies (cst), elaboration strategies (est), summary strategies (sst), and comprehension and remembering strategies (cr). On the other hand, memorizing strategies (mem) have a higher median for students in the first quartile of maths achievement, meaning that these students tend to rely more on this type of strategy than students in the fourth quartile of maths achievement.

Table 4 shows some descriptive statistics for the categorical student-level variables, where we also divided students in quartiles of achievement. This table shows that a greater percentage of students is female for all quartiles of achievement, except the fourth quartile, where 53.4% of the sampled students are male. Most students are not emigrant (imi), but the largest percentage of emigrant students happens for the first quartile (5.4%) and second quartiles (5.2%). Students in the first quartile of achievement attend mostly the third cycle grade (62.0%), whereas students in the fourth quartile attend mainly secondary education grade (91.6%). This means that students in the first quartile are likely to have repeated at least one year. In fact 69.7% of students in the first quartile have repeated at least one year (rep), while only 1.9% of students in the fourth quartile have repeated at least one year. Most students in our sample have a traditional family (fs), but in the first quartile of maths achievement a higher percentage of students (4.8%) have other type of families (this percentage is just 0.5% for students in the fourth quartile). About 40% of students in all quartiles have tutoring lessons outside the school (tut).

About 69% of students in the first quartile also have help from parents in their homework (hmw). Note, however, that for students in the fourth quartile, a big percentage (above 50%) also have support from their families in their homework. Regarding the expectations of students (exp), 88.9% of students in the fourth quartile expect to go to the university and take an undergraduate or postgraduate course, whereas in the first quartile only 26.4% of students have similar expectations. Most students in this quartile (44.1%) expect to finish just compulsory education (12 years) under professional courses (interestingly only 21% of these students indeed attend such type of courses).

Table 5 shows some descriptive statistics for the school-level continuous variables considered. There are 194 schools in our sample, 25% of which have less that 636 students and 25% have more than 1259 students. On average the Portuguese sampled schools have 8.74 students per teacher and 0.51 computers per student, where on average about 93% of these are connected to the internet.

Regarding descriptives for school-level categorical variables these are displayed in Table 6. 88.1% of the schools analysed are public, and only 3.1% are fully private. Note that private schools in Portugal may receive some funding from the government when they are located in a region without public schools. As a result these private schools are obliged to accept all students from that region. These schools are called private dependent in our sample and they are more alike public schools than private schools. This means that our sample of Portuguese schools is under represented of private schools (e.g. In the country there are about 80% of public secondary schools. In our PISA sample this percentage is 88%). Schools sampled are mainly located in villages (83.0% of schools, corresponding to 83.8% of students attending

Table 4: Descriptive statistics of student-level categorical variables per quartile of maths achievement

		Students in 1 quartile		Students in 2 quartile		Students in 3 quartile		Students in 4 quartile	
Variable	Values	Freq.	Perc.	Freq.	Perc.	Freq.	Perc.	Freq.	Perc.
sex	Female	588	61.9%	574	57.8%	559	53.5%	425	46.6%
	Male	362	38.1%	419	42.2%	486	46.5%	487	53.4%
imi	No	899	94.6%	941	94.8%	1004	96.1%	883	96.8%
	Yes	51	5.4%	52	5.2%	41	3.9%	29	3.2%
grade	Third cicle	589	62.0%	382	38.5%	187	17.9%	77	8.4%
	Sec. educ. (general)	164	17.3%	474	47.7%	754	72.1%	797	87.4%
	Sec. educ. (professional)	197	20.7%	137	13.8%	104	10.0%	38	4.2%
fs	Traditional Family	734	77.3%	811	81.7%	895	85.6%	773	84.8%
	Just one element	170	17.9%	159	16.0%	138	13.2%	134	14.7%
	Other	46	4.8%	23	2.3%	12	1.1%	5	0.5%
tut	No	534	56.2%	600	60.4%	620	59.3%	580	63.6%
	Yes	416	43.8%	393	39.6%	425	40.7%	332	36.4%
hmw	No	296	31.2%	377	38.0%	439	42.0%	426	46.7%
	Yes	654	68.8%	616	62.1%	576	58.0%	486	53.3%
exp	9 grade	166	17.5%	22	2.2%	4	0.4%	2	0.2%
	12 grade professional	419	44.1%	308	31.0%	112	18.4%	65	7.1%
	12 grade general	114	12.0%	128	12.9%	92	8.8%	34	3.7%
	under/postgraduate course	251	26.4%	535	53.9%	757	72.4%	811	88.9%
rep	No	288	30.3%	685	69.0%	961	92.0%	895	98.1%
	Yes	662	69.7%	308	31.0%	84	8.0%	17	1.9%

Table 5: Descriptive statistics of school-level continuous variables

	Mean	St Dev	Q_1	Median	Q_3
DIM	968.30	464.38	636.00	892.00	1259.00
PG	50.56	4.71	47.72	50.30	53.60
PTUT	0.40	0.14	0.31	0.42	0.48
STR	8.74	2.72	7.15	8.33	9.54
C/S	0.51	0.29	0.33	0.46	0.64
COMPWEB	0.93	0.17	0.96	1.00	1.00
STBEA	0.11	0.94	-0.50	0.05	0.65
TEABEA	0.19	0.91	-0.27	-0.03	0.71
EXTC	0.36	0.81	-0.25	0.32	0.96

schools in village areas). We consider villages those locations with less than 100 000 inhabitants.

Table 6: Descriptive statistics of school-level categorical variables

		Students		Schools	
Variable	Values	Frequency	Percentage	Frequency	Percentage
TYPE	Public	3449	88.4%	171	88.1%
	Private Dependent	337	8.7%	17	8.8%
	Private Independent	114	2.9%	6	3.1%
LOC	Village	3267	83.8%	161	83.0%
	City	633	16.2%	33	17.0%

4 Empirical results

The multilevel quantile regression model defined in (5) was applied to our sample of Portuguese students, where first only student-level variables were considered and then school level variables were accounted for (as explained in section 3). We approximated the log-likelihood using a Gauss-Hermite quadrature with K = 9 nodes for 5 quantiles $\tau \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$ and optimized the objective function via Nelder-Mead. Standard errors were computed using R = 50 bootstrap replications. In what follows we present and discuss the results from the multi-level quantile analysis, showing the final results of the models that included both student and school variables.

4.1 Determinants of student success in Portugal

The estimated regression coefficients and their standard errors, the estimated variance of the random effects, $(\hat{\psi}_u^\tau)^2$, the ICC and AIC are reported in Table 7. The least squares solution from a linear mixed-effects model (or multi-level linear model) is also presented for comparison. Note that a null model resulted in 30% of the total variability in math test scores being explained by school differences. This clearly justifies the use of multilevel models.

In Table 7 we show only the variables that revealed as statistically significant. Our first finding regards, therefore, the variables that were not considered important in explaining students results: at the student level, these were the family structure of the student and its home possessions (a variables that was included just for control as the escs variables already included home possessions of students). Regarding school level variables, in spite of the large variety of variables considered and the several attempts performed with different specifications, no variable revealed statistically significant in explaining the results of Portuguese students in math, except the location of the school.

The estimated coefficients for each model quantile in Table 7 are generally interpreted (in the case of continuous independent variables) as the estimated change on Math test scores, resulting from a unitary increase in each independent variable in turn, for groups of students in different rank positions. Such an interpretation cannot be used in our case given endogeneity problems of our estimated model (several endogeneity problems may happen in education contexts some relating to simultaneity of causality, as we will see below, but also with the fact that students are not randomly assigned to schools and the selection of a school is related with student characteristics). As a result we will interpret our coefficients as a correlation between the dependent and independent variables, and analyse whereas this correlation is

	Linear Mixed	model	$\tau = 0.1$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.75$		$\tau = 0.90$	
	Value	Std. Error	Value	Std. Error	Value	Std. Error	Value	Std. Error	Value	Std. Error	Value	Std. Error
Intercept	454.393*	6.005	392.341*	12.693	418.346*	12.630	455.877*	9.593	484.620*	8.80556	528.322*	17.652
sex (ref: Female)												
Male	30.762*	1.933	29.041*	1.798	30.776 *	2.713	31.835*	1.064	29.754*	1.73004	29.465*	1.899
escs	10.794*	0.960	11.476*	2.009	10.845*	1.182	12.388 *	0.818	10.680 *	1.121	9.691*	1.754
imi (ref: No)												
Yes	-15.805*	4.551	-13.325	5.928	-16.257	9.203	-13.957	5.552	-21.598*	6.981	-12.506	5.424
hmw (ref:No)												
Yes	-9.026 *	1.909	-8.380*	3.584	-8.699*	2.751	-10.240*	2.970	-9.086	3.474	-8.132	3.167
tut (ref: No)												
Yes	-11.061*	1.934	-12.634*	5.735	-9.460	3.925	-11.193*	2.833	-10.410*	1.972	-11.140 *	2.094
exp (ref: 9 grade)												
12 grade professional	42.617*	4.646	36.448*	8.955	42.512*	7.524	42.082*	8.420	47.029*	9.148	40.334*	13.812
12 grade general	32.839*	5.401	25.437*	8.147	28.362	11.559	33.686*	10.623	39.946*	10.199	33.399	16.746
Posgraduate course	54.252*	5.009	42.416*	8.964	52.544*	9.539	53.726*	11.022	62.325*	11.5392	55.9549*	18.6454
rep (ref:No)												
Yes	-51.842*	3.384	-52.606*	6.690	-50.994*	7.408	-49.793*	4.419	-50.092*	1.581	-58.463*	4.8952
grade (ref: Third cicle)												
secondary educ. (general)	14.024*	3.420	11.172*	8.262	10.237	8.136	15.967*	4.203	16.494*	3.145	14.188*	4.566
secondary educ. (professional)	-19.041*	3.836	-13.191*	5.982	-18.251*	5.416	-19.744*	7.003	-17.331	9.094	-26.729	13.437
cst	6.639*	1.420	8.869*	0.659	7.587*	1.143	5.027*	1.210	5.453*	1.556	2.393	1.37
est	5.171*	1.467	4.266*	1.867	4.108	2.575	5.557*	0.716	4.714	2.036	6.874*	2.181
mem	-11.161*	1.040	-11.293*	2.364	-13.119*	0.915	-10.676*	0.956	-9.89 *	1.377	-9.235*	1.425
sst	8.626*	1.119	9.535*	2.292	9.114*	1.806	9.507*	1.260	8.817*	1.270	8.062*	2.421
cr	8.391*	1.049	8.279*	2.064	7.031*	0.666	8.461*	1.164	7.980*	1.749	8.714*	2.674
loc (ref: Village)												
City	12.334*	4.510	20.985*	7.239	13.074*	2.944	13.027*	2.723	14.008	6.222	14.028*	2.998
$(\hat{\psi}_u^\tau)^2$	337.824		681.215		438.762		387.499		433.562		634.037	
ICC	0.096		0.070		0.074		0.090		0.074		0.066	
AIC	42675		44469.01		43518.33		43042.63		43470.85		44438.08	

* statistically significant at the 5% level

Table 7: Estimated coefficients for multilevel quantile regression model

different depending on the percentage of students lying below the estimated regression line (corresponding to the quantile).

We will analyse the coefficients through a graphical analysis where variables coefficients (y-axis) are plotted for each quantile model (x-axis).

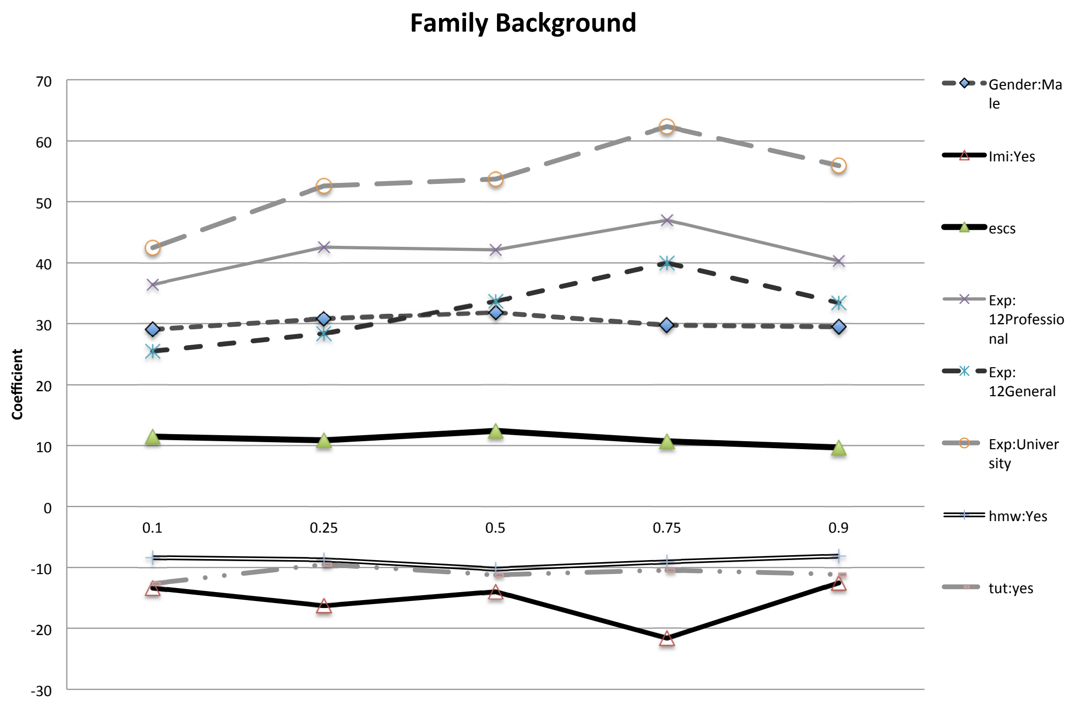

Figure 1: Estimated Coefficients - demographic and family factors

In Figure 1 we can see that as far as family factors are concerned, family expectations (exp) play an important positive role in determining student success. When family expectations are high (university courses) students have on average more 40 points in PISA scores (if they are in quantile 10%) or more 60 points (if they are in quantile 75%) relative to low family expectations (third cycle). Note that in all cases the coefficients rise with the quantile of achievement of students meaning that expectations have an higher correlation with the scores of high quantile achievers than of low quantile achievers. The association of escs to math performance is positive (about 10 points) and about constant per quantile, suggesting that one should expect about the same impact of socio-economic background on students achievements irrespective of the rank of the student in the math scores' distribution. Gender also shows a stable and positive association with math scores (with boys scoring about 30 points above girls). A negative correlation with math scores is observed for the immigration situation of the student.

Interestingly family support in homework shows a negative coefficient in Math test scores for all quantiles, and this coefficient is very close to the coefficient associated to private tutoring outside the school, making one believe that the two variables may be capturing a similar effect. For both cases the coefficients do not vary much per quantile of achievement, suggesting that the penalising effect of family support and private tutoring may not depend on the rank position of the student. These findings are somehow surprising - first the negative effect is unexpected and secondly one would expect that external help benefited more students in lowest positions in the overall rank of achievement. The investigation of the reasons behind this surprising finding are outside the scope of the present article. We believe that some endogeneity problems may cause the observed behaviour of these variables (relating to the simultaneity of causality), meaning that the model cannot estimate the exact impact of help in homework or outside tutoring lessons. We attempted some possible forms of solving the problem, like the inclusion of student level variables averaged per school (as shown in Hanchane and Mostafa (2012)) and also attempted to find some instrumental variables like home possessions that being a measure of economic status of the student may impact more strongly the attendance of tutoring lessons outside the school, than performance per se. Since, the problem remained for every specification attempted we leave this subject as a topic for future research.

Regarding cognitive and learning variables, these are shown in Figure 2.

The variable that indicates whether the student has already repeated one or more years (rep) has a negative relation with maths test scores (repeating students obtain on average about 50 points less in maths PISA scores than students that never repeated in all ranks of achievement except for the 90% quantile where the loss is almost 60 points). The coefficients associated to the grade of the student show advantage for students in secondary education and disadvantage for students in professional courses, in relation to students in the third cycle. For the students in secondary education lower coefficients are visible for students in low quantiles of achievement, and for students in professional courses lower coefficients are visible for students in higher quantiles of achievement. This means that attendance of professional courses

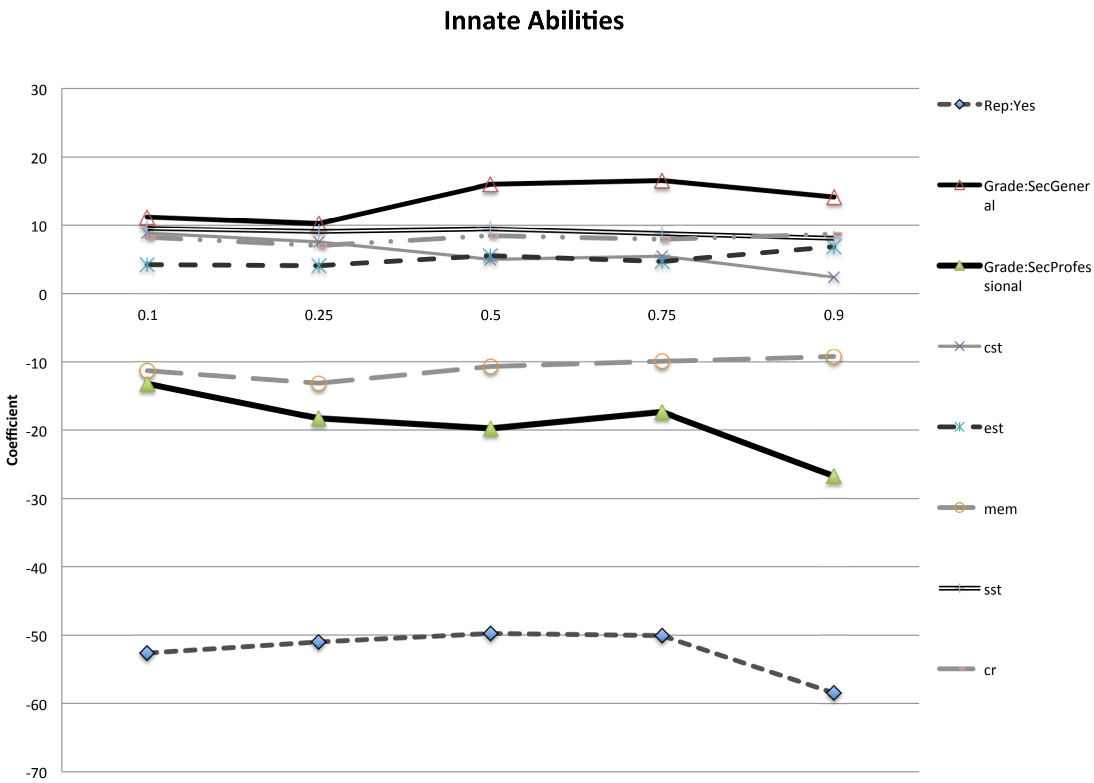

Figure 2: Estimated Coefficients - cognitive student factors

penalises more students in the highest quantiles of achievement and the attendance of secondary courses benefits less the students in lower quantiles of achievement.

As far as study strategies are concerned, we observe that they all have a positive relation to Math scores, except for memorization strategies that have a negative correlation. In most cases the coefficients are constant across quantiles, except for control strategies (cst) which show the lowest impact for students in the highest rank positions. Recall that these strategies are related to the extent to which students control their learning process (they understand what they need to learn, they figure out what they did not understand, and they look for additional information) and these kind of strategies may benefit most the worst students.

Perhaps the most surprising finding of this paper is the unimportance of school variables (and all of those in Table 1 were attempted). Only one school level variable appeared to be relevant in explaining the observed school differences: the location of the school (city or village). Schools located in cities tend to have a positive impact on students outcomes making them obtain, on average, about 10 points more than students in village schools. Interestingly the beneficial effect of cities is about the double for students in the lowest quantiles of achievement.

Note that many previous studies have reported the relatively small importance of school-level variables in explaining differences between schools. For example, Hanushek (1986) reviewed a number of studies in the literature and concluded that out of 65 studies analysing expenditure per pupil as a determinant of success, only 13 studies revealed a statistically significant and positive relationship. Teacher related variables, like teacher education or experience, also showed the positive and statistically significant expected effect in a reduced number of cases. In Woessmann (2003) the authors even report a negative influence of expenditure per student (at the country level) on student's achievement, and a negative impact of smaller class sizes.

Clearly, it is not the case that teachers are unimportant or that class size does not matter. The case is that the proxies used to measure the importance of teachers on schooling, and the importance of other school variables, may be poor proxies of what they are intended to capture.This is probably because school effects are mainly related with the quality of the teachers at the school, their capacity to motivate students and to make them overcome their limitations. However, quantitative measures to account for these factors are yet to be disclosed.

Finally a note regarding the quality of the fit of the estimated models. AIC values show that the quality of fit is very similar across all estimated models. When the multilevel linear model is used, inter-schools variability (measured by ICC) accounts for 9.6% of the total variation found in the maths test scores. It is interesting to note that when the quantile regression is used the ICC follow an inverted u-shape curve, with schools explaining less variability in test scores for students in the lower and upper quantiles, and explaining more for students in the 50% quantile. This behaviour corroborates the finding of Geraci and Bottai (2014) that differences between schools play a less important role in explaining variability in test scores of the lowest performers and the best performers. This finding is intuitive for the best performing students, where their grades may be more a result of their innate abilities and effort rather than of school's efforts. However, for the lowest performing students it was expected that schools played a more important role in explaining variability of results amongst students. The reason for this not happening, may be related to the fact that given heterogeneity of students in classes, teachers tend to focus in general on the mid achieving student. This means that differences between schools will, as a result, be felt more for this type of students. Further analysis of the Portuguese and other contexts should aim at corroborating (or not) this finding.

4.2 The school effect and its determinants

In the previous section our focus was to understand which student level and school level variables could explainin Maths achievement in different quantiles. In this section we estimate individual school effects (the estimate of u_{0j}) based on the procedure of Geraci and Bottai (2014), in order to understand the degree of the heterogeneity of Portuguese schools identified in the previous section. In quantile regression, most of the errors for lower quantiles are positive (most students lie above the 10% quantile regression plane), and most errors are negative for higher quantiles (most students lie below the 90% quantile regression plane). The procedure used in Geraci and Bottai (2014) allows one to transform all the error terms to be around zero, such that their interpretation is similar to that of non-quantile models. These effects are interpreted as the difference between the scores of pupils attending a particular school and the scores of all pupils analysed. When the school effect is positive, it means that the school achieves higher than expected grades; when it is negative it means that the school scores are below expected.

Figure 3 shows the global picture of average school effects obtained for each quantile regression, and also the school effects of four selected schools.

On average school effects are around zero, but negative for the quantiles 10% and 25% and positive for the quantiles 75% and 90%. This means that on average schools tend to show a performance lower than expected for students in low ranks of achievement and above expected for students in higher ranks of achievement. It is interesting to note that there are huge differences between the best performing and the worst performing schools in our sample. School 73 is amongst the best performers and for all quantiles students of this school show an average attainment about 80 points higher than that of students attending the lowest performing school (school 196 in Figure 3). All schools in our sample show very similar profiles in terms of their effects over different quantiles of students, where in general a growing trend is observed from quantile 10% to quantile 90% in terms of school effects. In spite of the general behaviour observed in Figure 3, some schools show different profiles of effects. For example, school 83 shows an inflection for students

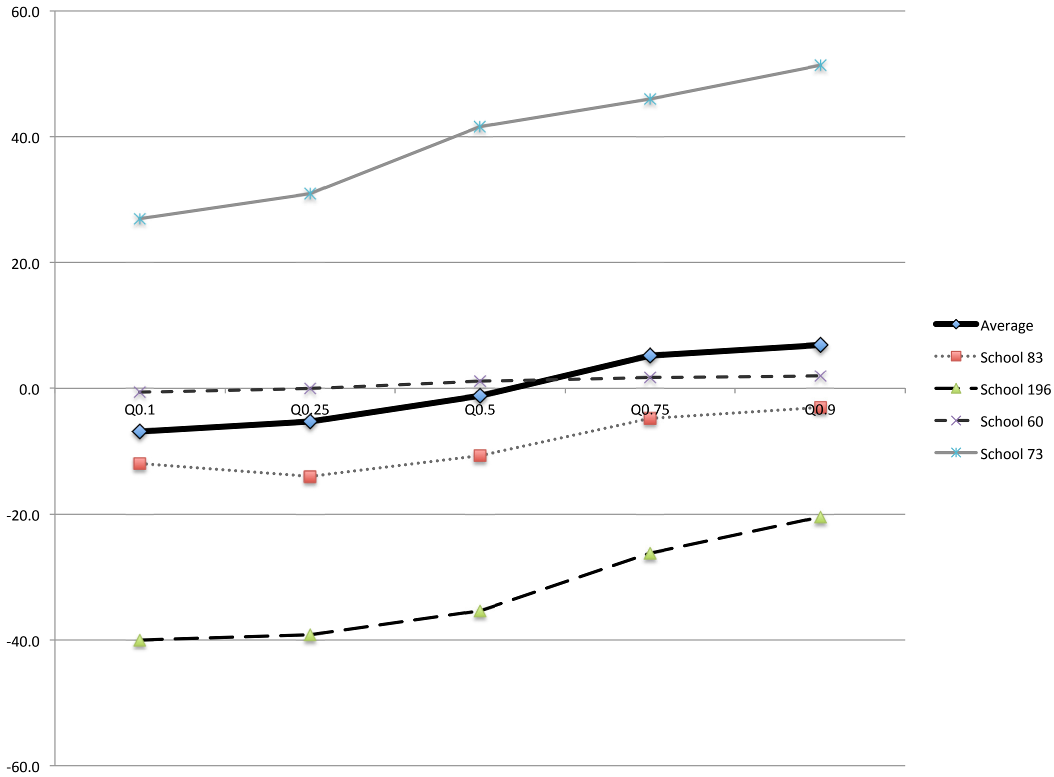

Figure 3: Estimated School Effects

in the lowest quantile, where the school has a better effect for students in quantile 10% than for students in quantile 25%. School 60, on the contrary exhibits a very stable effect on students over all quantiles of achievement.

These findings corroborate previous analysis in Portuguese schools, where schools were found to be differentially effective - showing a stronger effect on good students and a lower effect on less able students. This was more evident in the case of secondary education. For the third cycle of education the same was found, but there was an inflection in the trend for the lowest achievers similar to what we observed above for school 83 (see e.g. Portela (2014), who reported on an analysis of the VA of schools participating in an external evaluation programme in Portugal, or Romão (2012)). Given that most of Pisa sampled students attend secondary education, the findings from this paper corroborate our previous findings that school efforts tend to benefit more the students in higher quantiles of achievement than lower performers.

5 Conclusion

This paper analysed the performance in PISA-2009 Math test scores of Portuguese students through a multilevel quantile regression model. The objective of the analysis was twofold. In one hand we wished to understand the drivers of students' success in maths and whether the impact of these drivers differed for students located in different positions of the ranking of test scores; and on the other hand, we wished to understand the magnitude of school effects, and the extend to which these effects could be explained by some school level variables.

The main results of our analysis point to the existence of some stable and important effects of gender, socio-economic background, and repetition, and important and growing (with the rank position of the student) effects of the grade attended (advantage for students attending the secondary education), and the students' expectations. Interestingly study strategies appeared as important variables in explaining achievements, with most of them showing stable impacts across different students' rank in maths, except for control strategies that proved more beneficial for students with the lowest math scores. Two variables showed counter-intuitive signs (the existence of tutoring classes outside the school and the help in homework by parents). The negative impact of these variables is likely to be related with endogeneity problems, since the negative sign identifies a negative relationship between the variables and not a cause-effect relationship. Future work should aim at correcting these problems.

From our analysis it became apparent that schools still played an important role in explaining students' success in maths even after student-level variables being accounted for. The highest impact of schools happened for students in the 50% quantile, with schools explaining about 9% of the variability in test scores. The lowest impact of schools happened in the extreme quantiles meaning that for very good or very 'bad' students the school attended does not seem to make as much difference as for students in the middle of the distribution.

We also show in this paper that Portuguese schools can have very different effects on student's achievement, with schools showing math scores above expected in all quantiles considered, and others showing math scores below expected for all the quantiles. Interestingly school effects are on average increasing with the student's ranking. This means that in general students in the highest ranks are the ones that show the smallest distances from expected, and students in the lowest ranks are the ones showing the largest distances from expected. This is coherent with some previous studies in the portuguese context, showing that portuguese schools tend to foster and improve the achievements of its best students at the expense of the low achieving students that may be left behind. Finally note that school level variables were mostly non-significant in explaining students achievement. Only the location appeared to justify students test scores. It is our belief, that a number of other variables could contribute to explain the different effects of schools on students' achievement, but these are difficult to measure as they probably relate mainly to unmeasured teacher effects.

Acknowledgements

The authors acknowledge the financial support of the Portuguese Foundation for Science and Technology (FCT) through project PTDC/GES/68213/2006.This research was financed by Portuguese Funds through FCT - Fundação para a Ciência e a Tecnologia", within the Project UID/MAT/00013/2013. The authors are also grateful to Ricardo Ribeiro for the useful comments on the paper.

References

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

Agasisti, T., Ieva, F., and Paganoni, A. (2014). Heterogeneity, school effects and achievement gaps across italian regions: further evidence from statistical modeling. MOX report N. 7/2014.

Alacaci, C. and Erbas, A. K. (2010). Unpacking the inequality among turkish schools: Findings from pisa 2006. *International Journal of Educational Development*, 30:182–192.

Aristovnik, A. and Obadic, A. (2014). Measuring relative efficiency of secondary education in selected EU and OECD countries: the case of slovenia and croatia. *Technological and Economic Development of Economy*, 20(3):419–433.

Berberoglu, G. and Tansel, A. (2014). Does private tutoring increase student academic performance? evidence from turkey. *International Review of Education*, (60):683–701.

Carneiro, P. (2008). Equality of opportunity and educational achievement in portugal. *Portuguese Economic Journal*, 7(1):17–41.

Costanzo, A. (2015). The effect of m@tabel on italian students performances: A quantile regression approach. *Procedia - Social and Behavioral Sciences*, 197:236–244.

Dias, V. and Ferrão, M. (2006). Modelo multinivel do desempenho escolar de alunos socialmente desfavorecidos em escolas públicas/privadas - aplicação aos dados portugueses do pisa 2000. *Psicologia e Educação*, 2:63–162.

Ferrão, M. (2009). Sensivity of value added model specifications: measuring socio-economic status. *Revista de Educacion*, 348:137–152.

Ferrão, M. and Goldstein, H. (2009). Adjusting for measurement error in the value added model: evidence from Portugal. *Quality and Quantity*, 43:951–963.

Fonseca, J., Valente, M., and Conboy, J. (2011). Student characteristics and pisa science performance: Portugal in cross-national comparison. *Procedia Social and Behavioral Sciences*, 12:322–329.

Fuchs, T. and Woessmann, L. (2004). What accounts for international differences in student performance? a re-examination using PISA data. IZA Discussion paper series, No. 1287.

Geraci, M. (2014). lqmm: Linear quantile mixed models. *R package version 1.5.*

Geraci, M. and Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric laplace distribution. *Biostatistics*, 8(3):140–154.

Geraci, M. and Bottai, M. (2014). Linear quantile mixed models. *Statistics and Computing*, 24(3):461–479.

Giambona, F. and Porcu, M. (2015). Student background determinants of reading achievement in italy. a quantile regression analysis. *International Journal of Education Development*, 44:95–107.

Goldstein, H., Rasbash, J., M., Y., and G., W. (1993). A multilevel analysis of school examination results. *Oxford Review of Education*, 19:425–433.

Gray, J. Goldstein, H. and Jesson, D. (1996). Changes and improvements in school's effectiveness: trends over five years. *Research Papers in Education*, 11(1):35–51.

Hanchane, S. and Mostafa, T. (2012). Solving endogeneity problems in multilevel estimation: an example using education production functions. *Journal of Applied Statistics*, 39(5):1101–1114.

Hanushek, E. (1979). Conceptual and empirical issues in the estimation of educational production functions. *The Journal of Human Resources*, 14(3):351–388.

Hanushek, E. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, XXIV:1141–1177.

Hanushek, E. and Taylor, L. (1990). Alternative assessments of performance of schools: Measurement of state variation in achievement. *The Journal of Human Resources*, 25(2):179–201.

Hanushek, E. A., Link, S., and Woessmann, L. (2013). Does school autonomy make sense everywhere. *Journal of Development Economics*, 104:212–232.

Koenker, R. and Bassett, G. (1878). Regression quantiles. *Econometrica*, 46:33–50.

Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles. *Journal of Econometrics*, 128(1):137–164.

Kruger, N. (2011). The segmentation of the argentine education system: evidence from pisa 2009. *Regional and Sectoral Economic Studies*, 11(3):41–64.

Ladd, H. F. and Walsh, R. P. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, 21(1):1–17.

Mancebón, M., Calero, J., Choi, A., and Ximénez-deEmbún, D. (2012). The efficiency of public and publicly subsidized high schools in spain: Evidence from pisa-2006. *Journal of the Operational Research Society*, 63:1516–1533.

Martins, L. and Veiga, P. (2010). Do inequalities in parents' education play an important role in pisa students' mathematics achievement test score disparities? *Economics of Education Review*, 29:1016–1033.

O'Donoghue, C., Thomas, S., Goldstein, H., and Knight, T. (1997). 1996 DfEE study of value added for 16-18 year olds in England. *DfEE Research Series*, March.

OECD (2008). Measuring improvements in learning outcomes: Best practices to assess the value-added of schools. *OECD Publishing*.

OECD (2010a). Pisa 2009 results: Learning to learn - student engagement, strategies and practices (volume iii). *OECD Publishing*, (http://dx.doi.org/10.1787/9789264083943-en).

OECD (2010b). Pisa 2009 results: What makes a school successful? resources, policies and practices (volume iv). *OECD Publishing*, (http://dx.doi.org/10.1787/9789264091559-en).

OECD (2012). Education at a glance 2012. *OECD Publishing*.

Pereira, M. (2011). Uma análise da evolução do desempenho dos estudantes portugueses no programme for international student assessment (pisa) da ocde. *Boletim Económico, Banco de Portugal*, pages 131–146.

Pereira, M. and Reis, H. (2014). Retenção escolar no ensino básico em portugal: determinantes e impacto no desempenho dos estudantes. *Boletim Económico, Banco de Portugal*, pages 63–87.

Perelman, S. and Santin, D. (2011). Measuring educational efficiency at student level with parametric stochastic distance functions: an application to spanish pisa results. *Education Economics*, 19(1):29–49.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2014). nmle: Linear and nonlinear mixed effects models. R package version 3.1-118.

Portela, M. (2014). Efficiency and value-added in secondary education: The case of portugal. Paper presented are the workshop Efficiency in Education, September, 2014, London, UK.

Raudenbush, S. and Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20:307–335.

Romão, P. (2012). Alguns fatores determinantes dos resultados obtidos pelos alunos do 9 e 12 anos nos exames nacionais de português e matemática e o efeito escola. PhD thesis, Faculdade da Educação e psicologia da Universidade Católica Portuguesa.

Suárez-Álvarez, J., Fern'andez-Alonso, R., and niz, J. M. (2014). Self-concept, motivation, expectations, and socioeconomic level as predictors of academic performance in mathematics. *Learning and Individual Differences*, 30:118–123.

Tian, M. and Chen, G. (2006). Hierarchical linear regression models for conditional quantiles. *Science in China: Series A Mathematics*, 49:1800–1815.

Valente, M., Fonseca, J., and Conboy, J. (2011). Inquiry science teaching in portugal and some other countries as measured by pisa 2006. *Procedia Social and Behavioral Sciences*, 12:255–262.

Woessmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2):117–170.

Yu, K. and Zhang, J. (2005). A three-parameter asymmetric laplace distribution and its extension. *Communications in Statistics, Theory and Methods*, 34:1867–1879.