Artigo submetido a 1 de Outubro de 2019; versão final aceite a 10 de Abril de 2020
Paper submitted on October 1, 2019; final version accepted on April 10, 2020

Territorial Differences in Student Performance in Portugal: The Role of Family Characteristics and School Composition

Diferenças Territoriais no Desempenho dos Alunos em Portugal: O Papel das Características Familiares e da Composição Escolar

Alcino P. Couto
acouto@ubi.pt
University of Beira Interior and NECE Research Unit in Business Sciences/UBI

Patrícia Costa
patriciamotacosta@gmail.com
Centro de Matemática Aplicada à Previsão e Decisão Económica (CEMAPRE), European Commission, Joint Research Centre, Human Capital and Employment Unit

Maria Eugénia Ferrão
meferrao@ubi.pt
Universidade da Beira Interior, REM - Research in Economics and Mathematics (CEMAPRE)

Abstract/ Resumo

This research represents an early attempt to disentangle the role of families, schools and regions characteristics in student performance in order to support the design of educational-related policies. Multilevel statistical models applied to PISA 2015 Portuguese data are used. The findings reveal the predictive power of socioeconomic variables operating through family background and the school composition. The regional variance explained by individual or school composition concerning the socioeconomic status ranges from 20% to 24%. The impact relies mainly on variability across schools. The variation explained due to region differences is smaller than 5%. After controlling for socioeconomic status, some regions present lower regional variance than the metropolitan region of Lisboa, showing that their students tend to achieve higher scores. Some research and policy implications are discussed. The regional dimension allows expanding the avenues to exam-

Esta pesquisa representa uma tentativa inicial de dissociar o papel das características das famílias, escolas e regiões no desempenho dos alunos, com vista a apoiar a formulação de políticas educacionais. São utilizados modelos estatísticos multinível para os dados portugueses do PISA 2015. Os resultados revelam o poder preditivo das variáveis socioeconómicas que influenciam o desempenho dos alunos associadas ao contexto familiar e à composição escolar. A variância regional explicada pela composição individual ou escolar em relação ao nível socioeconómico varia de 20% a 24%. O impacto depende principalmente da variabilidade entre as escolas. A variação explicada devido às diferenças regionais é menor que 5%. Controlado o nível socioeconómico, algumas regiões apresentam menor variância regional do que a região metropolitana de Lisboa, mostrando que os seus alunos tendem a obter pontuações mais elevadas. São discutidas algumas implicações de investigação e de politi-

channels through which educational resources exert their impact on student performance. In turn, the decentralisation of educational policies seems to be required to pursuit suited local educational governance structures to address the effects of disadvantaged backgrounds.

Keywords: family background; school composition; regional achievement-gaps; educational resources; socioeconomic status and educational-related policies

JEL Code: I24, I260

cas. A dimensão regional permite ampliar as formas de examinar quais os recursos educacionais que influenciam o desempenho dos alunos. Por sua vez, a descentralização das políticas educacionais parece ser necessária na busca de estruturas adequadas à governação educacional local para lidar com os efeitos em grupos desfavorecidos.

Palavras chave: contexto familiar; composição escolar; diferenças regionais de desempenho; recursos educacionais; nível socioeconómico e políticas relacionadas com a educação

Código JEL: I24, I260

1. INTRODUCTION

A growing body of empirical evidence on student performance acknowledges that the national educational landscapes tends to present remarkable territorial variations in educational attainment indicators, educational resources and learning opportunities. This emerging stream of educational literature goes further, sustaining that to target educational inequalities research and educational policies need to take into account the determinants of regional variations concerning the level and inequality of students' performance (see e.g. SMF, 2016; Carnoy et al., 2015; Agasisti & Cordero-Ferrera, 2013; Agasisti & Vittadini, 2012; Pereira & Reis, 2012; Donato & Ferrer-Esteban, 2012; Bratti, Checchi & Filipin, 2007; Woessmann, 2007 and Brasington, 2002).

Based on a survey of selected empirical studies, Haahr et al. (2005) report that nine tenths of variation in student's performance takes place within countries. Carnoy et al. (2015) sustain that cross-country studies provide us with limited insights on the factors behind distribution of students' outcomes across countries. The authors suggest that differences in regulation of education systems, schools' systems and teacher labor markets are considered the major sources of variance in students' test scores across countries. Additionaly, other empirical evidences point out that both centralised and decentralised educational systems, namely in terms of teacher hiring and pay and curricula design, show significant regional educational performance inequalities (see. e.g. Agasisti & Cordero-Ferrera, 2013, Pereira & Reis, 2012; Donato & Ferrer-Esteban, 2012 and Bratti et al., 2007). Indeed, the effectiveness of international comparisons of educational systems is being questioned and the regional education comparative studies seem to offer new opportunities to a better understanding of the links between resources and outcomes associated to schooling processes.

One can observe that the educational literature provides two different perspectives on regional distribution of educational results and inequalities. The first one addresses the influence of regional socioeconomic disadvantages or advantages on educational outcomes and schooling process. The studies attempt to clarify to what extent the family background and school characteristics, as critical sources of educational resources, reproduce the existing inequalities between regions (see e.g. Agasisti & Vittadini, 2012; Pereira & Reis, 2012; Bratti et al., 2007; and Brasington, 2002). The second one focuses on the contribution of educational profile to the development and economic growth of regions. According to these premises, educational outcomes tend to reinforce existing economic and social inequalities between regions. (see e.g. Agasisti & Cordero-Ferrera, 2013, Checchi & Peragine, 2010, Rodrigues-Pose & Tselios, 2009, Shapiro, 2006, Arellano & Fullerton, 2005 and Berry & Glaeser, 2005).

In the case of Portuguese experience, we observe a lack of literature focused on regional analyses of PISA (The Programme for International Student Assessment) results. To the best

of our knowledge, there is no contribution beyond the research of Pereira & Reis, (2012), which is based on data from PISA 2009 (The Programme for International Student Assessment). Despite its centralised nature, the authors acknowledge that the Portuguese educational system reveals substantial heterogeneity regarding the level and inequality of student achievement scores. The indicators of educational achievement, of students and working-age population, show important territorial variation. This educational setting is entangled with a high level of socio-economic inequalities and a significant regional variation of human and educational capital (see Pereira & Reis, 2012 and Carneiro 2008). In spite of recent improvement in PISA ranking, educational and socio-economic inequalities indicators show that Portugal still ranks lower than the OECD average, particularly regarding the stock variables (see INE, 2018; OECD, 2018a and OECD 2018b).

It is well established in educational literature that differences in student achievement and educational performance signal the existence of efficiency and equity problems related with distribution of learning opportunities (see. e.g. OECD, 2016, Hanushek, 2016 and Reardon, 2012). High level of educational equity implies a combination of a high level of school efficiency and effectiveness and a marginal or neutral influence of the family background and other sources of regional heterogeneity. The exam of regional educational gaps, and their potential drivers, shed light on the sources of educational inequalities. Its observation and descriptive analysis are a first step to address the nature and the importance of regional educational inequalities.

Based on PISA 2015 data and using a multilevel model, this paper seeks to examine the regional distribution of student achievement in Portugal. This is an exploratory approach on the existence of regional differences in student achievement scores and whether they are associated with regional distribution of family and public educational resources. By addressing the role of family and school variables in shaping inequality within and across regions, we seek to identify new research questions and areas that educational policies need focus on.

Therefore, this study takes into account one measure of educational achievement: standardised test scores, controlling for sociodemographic features of students and family background variables. The main research questions are:

Q1- What is the percentage of variability of school results / success due to territorial differences?

Q2 - What is the percentage of variability in school results / success due to differences across schools?

Q3 - How much of this variability is due to socioeconomic factors?

The remainder of the paper is organised as follows. The next section addresses previous contributions to the debate on the drivers of educational regional heterogeneity. Section three presents both the multilevel models and data, and describe the results of statistical analysis. Section five reexamines the most important empirical evidences and reflects on their implications of educational policies. Section six presents final remarks.

2. IN SEARCHING FOR DRIVERS OF REGIONAL HETEROGENEITY

Hanushek (2016:3) argues that "The regional (…) disparities of today in education inputs are probably quite similar to those Coleman reported in 1966". This statement poses the relevant question whether ignoring territorial distribution of educational resources associated to families, schools and regions significantly distorts the explanation how educational resources and educational outcomes are correlated.

Overall, the regional focus on educational performance seeks to find out potential differences in regional education processes to examine drivers and how they are associated with students' achievements. In order to approach such research issues, variables associated with regional backgrounds are integrated as a component of the schooling process alongside with family backgrounds and school characteristics.

A significant part of empirical studies addressing regional variations of educational performance is being stimulated by PISA periodical reports, which landscape incorporates a large number of countries. Consequently, the studies address a high level of heterogeneity concerning political, institutional, socioeconomic and cultural settings that frame the schooling processes at subnational levels. New levels of education hierarchy and their attributes are addressed and many other aspects of

student, family and school enviroments could be explored.

Given the centralised profile of Portuguese educational system, Pereira & Reis (2012) examine the sources of regional heterogeneity of students' achievement. Based on PISA 2009 database, the study uses a standard education production function approach, considering as explanatory variables students' characteristics, family background and school resources. They attempt to quantify the impact on students' achievement of the student, family and school variables and determining what remains of regional variance after these variables are controlled for. The findings suggest that family backgrounds and school variables play an important role in explaining regional differences, albeit such differences appear to be more closely related to families rather than schools.

In order to estimate the impact of regional background - the *regional-effect*, the authors regress the remaining of regional variance on regional environment-related variables such as GDP per capita, regional development index, teachers experience in the regions, doctors per habitant, drop-out and literacy rates, divorce rate and crime rate. The results indicate a low predictive power of the *regional-effect*.

Taking into consideration the existing institutional differences between Spanish and Italian educational systems in terms of the level of decentralization, the similarities in terms of wealth, levels of educational achievement and the economic inequalities among regions, Agasisti and Cordero-Ferrera (2013) attempt to explain an observed *common feature* in the both countries: the existence of significant regional-achievement gaps across regions. Using PISA 2006 database and applying a multilevel analysis, the authors seek to explore potential differences in the regional educational production processes to explain variations in regional profile of educational students' achievement scores.

Although Spain has a decentralised government system, the findings do not support the assumption that regional variables exert a more relevant role in Spain than in Italy. On the contrary, the evidences suggest a more prominent role of the regional background in Italy setting. Indeed, the results indicate that the high heterogeneity across regions in both countries seems to be associated to different factors. While in Italy most of the regional variance is explained by differences between-schools (around 40%,) and regional backgrounds (10–12%), regional-achievement gaps in Spain are more correlated with difference within-schools (80-85%).

In the case of Italy, Agasisti and Cordero-Ferrera research (2013) shows that most of the regional variance and of the divergence between schools appear to be correlated with the tracking system (school-type differences), school resources, the role of the public investment and income inequalities across regions. The evidences from Spain demonstrate the relevance of differences within-schools, which suggest the potential influence of student and family characteristics on achievement scores. Indeed, the role of family background is appointed as a common feature, namely in what concerns immigrant status and cultural capital, even if the impact of these variables could be considered statistically modest. But these results need a careful interpretation because it appears to exist room to the role of family background effects in both cases. For example, in Italian case, the variance between schools is closely associated to tracking system. The literature on educational tracking demonstrates that tracking could increase variance in educational achievement, as well as that the extent of family background influence on student achievement tends to be statistically significant when schooling tracking takes place at an earlier age (see e.g. Woessmann, 2009 and 2007; Brunello and Checchi, 2007 and Hanushek and Woessman, 2006). In turn, the links between the divergences within-schools and home educational resources in Spain seems to be statistically significant.

Woessmann (2007) explores the analysis of territorial variation of educational performance in the context of the German federal state. The relevance of the investigation relies on the existence of substantial gaps in the level and inequality of student performance, as well as in the institutional structures of the school systems across states**.** Germany presents an institutional framework traced by a decentralised policy decision making process. Along with the existence of a national identity, a common language, a common legal system and a set of other common institutions, the states and local authorities share a high level of territorial and sectoral policy competencies that support different policy options.

The focus of Woessmann is the analysis of the correlations between the institutional features and the efficiency and equity in educational outcomes across the sixteen states. The

author operates an education production function approach and combines PISA 2003 data with variables from other sources. From the point of view of the efficiency, the findings positively correlate student performance with national exams, private school operation and socio-economic background. The resource endowment shows a negative association with student performance. With regard to the equality of opportunity the evidences reveal a robust association with delayed and less intensive tracking. This association is strongly anchored on the dependence of student achievement on their socio-economic background. Although no signs of efficiency-equity tradeoffs in educational production are observed across states, Woessmann (2007) points out that socio-economic background strongly impacts on student outcomes across German states, in line with the literature on education production functions. However, the evidences reveal that the states with higher average performance present lower levels of inequality, and both efficiency and equity are only achieved by private schools. These results pose the question concerning potential and unobserved linkages between family, school and region characteristics and their effects in predicting student achievement.

Despite the regional empirical studies previously reported to be associated with substantial differences in institutional framerworks, the most important findings suggest a similar path: the relevance of the predictive power of family and school backgrounds. Additionaly, we find other studies exploring regional education function production devoted to the examination of the factors behind the regional variance of educational outcomes. Likewise, their results appear also to support the observed trend with regard the role of the family, school and regional drivers.

Using data drawn from PISA 2006, Tomul and Celik (2009) explore regional variance across regions in Turkey. They attempt to identify key explanatory factors of the low levels of effectiveness and equity demonstrated by the national education system. As relevant findings, the authors point out, on the one hand, the impact of the family variables on students' achievement, even though their influence vary between the three knowledge fields analysed by PISA studies. On the other hand, taking into consideration the regional environment-related variables, the results show that the predictive power of family variables is positively correlated with the regional development level. The disadvantaged regions seem to be strongly associated to low level of student's proficiency. This linkage calls into question the role of the interactions between family characteristics and regional-environment related variables and the channels through which they operate. The research of Rodríguez-Pose & Tselios (2009) on education and income inequality in the regions of the European Union could provide us with some insights regarding the links between the two structures. Using the European Community Household Panel dataset for 102 regions over the period 1995–2000 and different static and dynamic panel data analyses, the authors argue that high levels of inequality in educational attainment are associated with higher income inequality. According to the research, the centrality of equity in both public policies agenda and societal engagement seems to promote a more effectiveness and equity educational systems, given the higher overall educational performance demonstrate by social-democratic welfare states, Protestant areas, and regions with Nordic family structures.

Mixing data from PISA 2003 and Italian territorial statistical sources, Bratti et al. (2007) examine the regional variance in students' mathematical scores. Considering a multilevel analysis based on family background, school types, resources and territorial features related to labour market, cultural resources and aspirations, the findings suggest a significant impact of buildings maintenance and employment probabilities. When regressing for territorial differences, the large part of North-South divide explanation is due to differences in resource endowments (75%). The school background account for the remaining of the regional variance. Continuing to consider Italian educational system as case study, Ferrer-Esteban (2011) attempt to examine in what extent the territorial inequalities affect the trade-off between effectiveness and equity in education system. The author explores the study of the educational heterogeneity using hierarchical regression models to approach variable variances at each level, and a combination of different sources from international evaluation surveys. The research confirms the observed North-South differences as adverse effect on students' scores. The results also underline the evidence of social heterogeneity between classes within schools, social segregation of schools, and the rate of teachers in precarious employment as sources of

educational performance variations across regions. While Bratti et al. (2007) point out the role of the allocation of resources, Ferrer-Esteban (2011) emphasises the school composition effect generated by the allocation of students across schools and classrooms. Given the centralised nature of Italian education system, its higher level of institutional homogeneity seems to show increased difficulties to guarantee educational effectiveness, as well as equality of learning opportunities. An educational question arises from such evidences. How important is the relationships between the level of decentralisation of education systems and educational-related policies with the educational effectiveness and equity?

Considering the critical role of educational equity to improve the overall of the efficiency of Spain educational system, Sicilia & Simancas (2018) examine, using data from PISA 2015, the correlation between socioeconomic and cultural variables and high educational performance. Taking into account the *Comunidades Autónomas* as territorial units, the authors analyse the achievement gaps observed between the lower and upper quartiles of the distribution of socioeconomic status, The Galicia exhibits a higher level of educational equity express by the lower level of the influence of socio economic variables. Castilla y León and the Basque Country follows Galicia. This set of regions presents a higher level of educational equity than the OECD and countries such as Finland, Canada or Norway. The findings shed light on the factors behind the educational trade-off between efficiency and equity. The most important question stems from the evidence that regions that stance in an intermediate level of development seem reveals a lower socioeconomic-achievement gap. The results provide room to a proative role of schools' influence.

Although the production of education seems to be similar across regions and significant regional variance of educational performance are identified, family and school backgrounds continue to play a critical predictive power. However, one can observe some differences how family and school characteristics are associated to student achievement. The studies analysed indicate that the relevance of their items appears to vary.

3. METHODOLOGY

3.1 Rationale for the study and statistical models

For the purpose of this study, we compare the results of two multilevel models: Model 1 (M1) is a two level model considering students (level one units) grouped in schools (level two units) with fixed effects for regions in the linear predictor; Model 2 (M2) the same as M1, with additive terms for student's socioeconomic level and socioeconomic school composition. In other words, M1 allows us to test the hypothesis of regional differencial fixed effects on student achievement in maths, science and reading, and the M2 equates such regional comparison for individual and school composition socioeconomic factors.

Functional specification of the statistical model for measuring student achievement in science, maths and reading, considering both the hierarchical structure and the regional location of schools, controlling for their individual and composition variables, requires separate models fitted for each cognitive domain. The total number of models fitted are six. Thus, we considered the two-level variance component model with pupils (indexed by i) at level 1, and schools (indexed by j) at level 2, where the student performance is the response variable (e.g. yScience). The model 2 we wish to estimate is written as follows:

$$yScience_{ij} \sim N(XB, \Omega)$$

$$yScience_{ij} = \beta_{0ij} + \sum_{p=1}^{24} \beta_p r_p + \beta_{25} SES_{ij} + \beta_{26} schoolSES_j \qquad \text{(eq. 1)}$$

$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$

$$u_{0j} \sim N(0, \sigma_{u0}^2)$$

$$e_{0ij} \sim N(0, \sigma^2_{e0}),$$

where the response variable $yScience_{ij}$ is the achievement in science of student *i* in school *j*. The explanatory variables r_p are dummy variables for every Portuguese region with the metropolitan region of Lisboa as baseline, the variable SES represents the student's socioeconomic status and it is measured by the ESCS[1] index, the variable schoolSES is the SES average per school and represents the socioeconomic school composition.

Equation (eq. 1) is specifyied for model two. The model one equation is similar but without the 25th and 26th additive terms. The term u_{0j} is the school-level random component, while e_{0ij} is pupil-level random error. Both random components are assumed to follow normal distributions.

3.2 Population and sample

The PISA addresses the extent to which students near the end of compulsory schooling have acquired the knowledge and skills that are essential for full participation in modern societies. This is a cross-sectional survey involving multiple-step sampling. The target population in each of the countries are the 15-year-old students who attend school between the 7th and 11th grades (OECD, 2014: 66). Sampling planning is stratified into two steps in which the primary sampling is school. Schools are selected with probability proportional to size. The second sampling unit is student. The sample size is 7325 students in 246 schools. The sample design was incorporated in the modeling through one of the procedures described in Pfeffermann et al. (1998) and implemented experimentally in MLwiN v2.31 (Rasbash, Browne, Healy, Cameron, & Charlron, 2014). The variables W_FSTUWT and W_FSCHWT were considered for sample expansion. For the assessment tests in Mathematics, Reading and Science, which are also applied in the PISA 2015 databases contain the results (plausible values) of the multiple imputation. For the purpose of this paper the variables were standardised for the Portuguese sample. The resulting scores have mean 0 and standard deviation approximately 1. For the remaining variables, descriptive statistics are presented in the appendix.

4. RESULTS

The models M1 and M2 presented above were fitted separately for cognitive domains of science, maths and reading. The variance partition coefficient based on the null model (without explanatory variables) shows that the variability across schools is approximately 30% and within schools is 70%. The Table 1 contains the results for M1 and in Table 2 for M2.

Regarding M1, the fixed parameters estimates suggest that the students who live in the regions of Açores and Tâmega e Sousa present on average lower achievements than their peers who live in Lisboa. This is observed for every cognitive domain assessed, meaning that students in those regions have their performance reduced more than half a standard deviation, i.e in Açores the reduction is on average -0.61 (sd=0.11) in science, -0.55 (sd=0.12) in maths, -0.56 (sd=0.11) in reading, and in Tâmega e Sousa, the reduction is -0.69 (sd=0.19) in science, -0.59 (0.20) in maths, and -0.72 (sd=0.19) in reading. In some regions, the performance reduction occurs just in one or two cognitive domains, which is the case of students attending schools located in the Alentejo Central, Baixo Alentejo and Alto Tâmega. In the remaining regions, there is no difference, statistically significant at the level of 5%, in comparison to the Área Metropolitana de Lisboa.

The results of M2 show that the socioeconomic level of the students' family and the school composition, have a strong predictive power of the students' scores in maths, reading and science. Depending on the domain, the explained variation ranges between 20% and 24% (cf. Table 3). In addition, we verified that, after controlling for socioeconomic differences at the individual and school level, there are regions that present on average higher performance than the Área Metropolitana de Lisboa. Somes examples are Área Metropolitana do Porto (0.20, s.d.=0.08 in science; 0.23 s.d.=0.08 in maths; 0.23, s.d.=0.08 in reading), Ave (0.33, s.d.=0.13 in science, 0.44 s.d.=0.14 in maths, 0.45 s.d.=0.15 in reading), Alentejo Litoral (0.39, s.d.=0.18 in maths), Beiras e Serra da Estrela (0.39, s.d.=0.18 in maths), Cávado (0.35, s.d.=0.14 in maths), r.a. Açores (0.21, s.d.=0.09 in maths; 0.21, s.d.=0.09 in reading), r.a.

[1] The PISA Index of Educational, Social and Cultural Status (ESCS) is a composite score derived from several indicators: parental education (PARED), highest parental occupation (HISEI), and home possessions (HOMEPOS). It also includes the number of books at home. It is built via principal component analysis (PCA). Higher values of ESCS indicate higher socio-economic status of the students.

Madeira (0.34, s.d.=0.11). This means that students attending schools located in such regions tend to achieve higher scores in at least one cognitive domain, than they would achieve if their schools were located in the Área Metropolitana de Lisboa.

Table 1. Estimates of model M1 parameters

Model M1	Science		Maths		Reading	
Fixed Part	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Constant	0.12	0.08	0.07	0.08	0.08	0.08
Alentejo Central	**-0.50**	**0.25**	**-0.51**	**0.26**	-0.37	0.29
Alentejo Litoral	-0.31	0.26	-0.24	0.27	-0.31	0.40
Algarve	-0.39	0.22	-0.32	0.22	-0.31	0.20
Alto Alentejo	-0.33	0.28	-0.23	0.29	**-0.35**	**0.18**
Alto Minho	-0.14	0.24	-0.10	0.24	-0.17	0.31
Alto Tâmega	**-0.90**	**0.33**	-0.61	0.33	**-0.92**	**0.35**
Área Metropolitana do Porto	-0.10	0.12	-0.08	0.13	-0.09	0.13
Ave	-0.36	0.20	-0.25	0.21	-0.25	0.23
Baixo Alentejo	-0.55	0.30	-0.48	0.31	**-0.52**	**0.23**
Beira Baixa	-0.36	0.28	-0.28	0.29	-0.32	0.31
Beiras e Serra da Estrela	-0.09	0.27	0.06	0.28	-0.18	0.14
Cávado	-0.37	0.20	-0.18	0.21	-0.23	0.22
Douro	-0.25	0.25	-0.28	0.26	-0.23	0.31
Lezíria do Tejo	-0.14	0.25	-0.23	0.26	-0.25	0.23
Médio Tejo	-0.32	0.25	-0.18	0.26	-0.28	0.33
Oeste	-0.31	0.23	-0.34	0.23	-0.32	0.30
R. A. Madeira	-0.43	0.23	-0.32	0.24	-0.33	0.29
R. A. Açores	**-0.61**	**0.11**	**-0.55**	**0.12**	**-0.56**	**0.11**
Região de Aveiro	-0.39	0.22	-0.31	0.22	-0.39	0.21
Região de Coimbra	-0.10	0.21	-0.03	0.21	-0.14	0.15
Região de Leiria	-0.18	0.23	-0.08	0.24	-0.18	0.27
Tâmega e Sousa	**-0.69**	**0.19**	**-0.59**	**0.20**	**-0.72**	**0.19**
Terras de Trás-os-Montes	-0.63	0.33	-0.50	0.34	**-0.74**	**0.21**
Viseu Dão Lafões	-0.22	0.23	-0.06	0.23	-0.16	0.25
Random Part						
School level	0.25	0.03	0.27	0.03	0.28	0.02
Student level	0.74	0.01	0.71	0.01	0.73	0.02
Number of units						
School level	246		246		246	
Student level	7325		7325		7325	

Table 2. Estimates of model M2 parameters

Model M2	Science		Maths		Reading	
Fixed Part	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Constant	**0.17**	**0.05**	**0.12**	**0.05**	**0.13**	**0.06**
Alentejo Central	-0.09	0.16	-0.12	0.17	0.04	0.24
Alentejo Litoral	0.31	0.17	**0.39**	**0.18**	0.33	0.23
Algarve	-0.03	0.14	0.03	0.14	0.04	0.13
Alto Alentejo	-0.06	0.18	0.04	0.19	-0.06	0.09
Alto Minho	0.11	0.15	0.15	0.16	0.06	0.26
Alto Tâmega	0.05	0.21	0.34	0.22	0.04	0.17
Área Metropolitana do Porto	**0.20**	**0.08**	**0.23**	**0.08**	**0.23**	**0.08**
Ave	**0.33**	**0.13**	**0.44**	**0.14**	**0.45**	**0.15**
Baixo Alentejo	-0.15	0.19	-0.06	0.20	-0.10	0.10
Beira Baixa	0.12	0.18	0.18	0.19	0.17	0.18
Beiras e Serra da Estrela	0.24	0.17	**0.39**	**0.18**	0.16	0.10
Cávado	0.17	0.13	**0.35**	**0.14**	0.31	0.13
Douro	0.15	0.16	0.12	0.17	0.18	0.15
Lezíria do Tejo	0.07	0.16	-0.03	0.16	-0.04	0.14
Médio Tejo	0.14	0.16	0.27	0.17	0.19	0.27
Oeste	0.28	0.15	0.27	0.15	0.28	0.22
R. A. Madeira	0.23	0.15	0.34	0.16	**0.34**	**0.11**
R. A. Açores	0.14	0.08	**0.21**	**0.09**	**0.21**	**0.09**
Região de Aveiro	0.03	0.14	0.11	0.15	0.04	0.13
Região de Coimbra	0.19	0.13	0.24	0.14	0.15	0.09
Região de Leiria	0.12	0.15	0.20	0.15	0.13	0.14
Tâmega e Sousa	0.14	0.13	0.24	0.13	0.14	0.14
Terras de Trás-os-Montes	-0.13	0.21	0.00	0.22	**-0.23**	**0.07**
Viseu Dão Lafões	0.05	0.14	0.21	0.15	0.12	0.12
ESCS	**0.21**	**0.01**	**0.23**	**0.01**	**0.19**	**0.01**
School_ESCS	**0.54**	**0.04**	**0.53**	**0.04**	**0.58**	**0.05**
Random Part						
School level	0.08	0.01	0.10	0.01	0.10	0.01
Student level	0.70	0.01	0.66	0.01	0.70	0.02
Number of units						
School level	246		246		246	
Student level	7225		7225		7225	

The Table 3 presents the decomposition of variance results for the six models fitted. The total amount of variance explained by M1 varies between 1% and 5% depending on the domain and it is null at the level of the student. The total amount of variance explained by M2 ranges from 20% to 24%, and is mainly due to the explained variability at the level of schools (23% in science, 21% in maths and 23% in reading). The explained variation increasing from M1 to M2 is due to the power of individual and school composition socioeconomic variables.

Table 3. Analyses of Variance

	Science		Maths		Reading	
	M1	M2	M1	M2	M1	M2
Variance explained, Total	1%	22%	2%	24%	5%	20%
Variance explained, School	6%	23%	5%	21%	5%	23%
Variance explained, Student	0%	4%	0%	5%	0%	3%

5. DISCUSSION

As the most important findings, the evidences suggest, on the one hand, the predictive power of socioeconomic variables operating through family background and the school composition. On the other hand, the limited influence of educational-regional environment. The results further seem to reinforce the resilience of regional-achievement gaps associated to differences in institutional and socioeconomic settings that frame the schooling processes at the subnational levels. The findings are in line with results consistently underscored by a large number of empirical studies on regional-achievement, namely previously mentioned. However, the interpretation of the results implies the consideration of methodological and modeling issues, as well as theoretical and empirical evidences.

With regard to methodological and modeling issues, some relevant aspects must be taked into account, given their potential bias effects. One of them is related with the cross-sectional approach. Such as the literature reports, most of what one can explore is the association between contextual factors and the outcomes based on the data collected at one point in time. Cross-sectional studies are not suited to approach causal inferences (see e.g. Caro, 2015). For instance, both the population mobility, steming from regional economic dynamics, and student's prior achievement issues are unlikely to be investigated by PISA 2015 data, since this is a crossectional survey. Another one limitation is associated with the application of contextualised results models (OECD, 2008). The comparison between contextualised results models and value added models show that the first ones tend to overquantify the estimates related to socioeconomic background variables (Ferrão, Barros, Bof, & Oliveira, 2018; Ferrão & Couto, 2013). The cautious interpretation of the findings is also justified by unobservable variables that not only exert influence on test scores, as well as they could be likely correlated with some of regressors. Moreover, regional variations in achievements could integrate the influence of unobservable *regional-effect* variables (see e.g. Pereira & Reis, 2012). This research does not consider all controlling variables related to regional environment, such as, for example, regional GDP per capita, population density and mobility, regional unemployment rate, occupational structure, and governance and regional educational policies. Finally, the schooling process is a complex one (see e.g. Pereira & Reis, 2012 and Buchmann, 2002). It is characterised by an intricate puzzle of links between family, school and regional characteristics (observable and unobservable). They interact and might be correlated with each other, giving raise to problems of endogeneity.

Notwithstanding the above caveats, the findings seem to corroborate one of the most important stylised facts of educational literature: the non-neutral role of the family background (see e.g. Hanushek, 2016, Reardon, 2012 and Buchmann, 2002). What is more, its influence is identified beyond the family home environment, in particular in school through socioeconomic school composition. However, the predictive power of family and school socioeconomic composition observed, highlighted by educational empirical studies across countries and regions, suggests reflexions on its evidences and educational policy implications.

Studies on the impact of the level of development on the predictive power of the family and school variables indicate that in disadvantaged socio-economic regions the influence of family variables on student achievement tends to be low and the impact of school characteristics much higher. Opposing signals are observed in advantaged regions (see e.g. Tomula & Celik 2009; OECD 2005; Woessmann, 2004 and Heyneman and Loxley, 1989).

However, the findings indicate, after controlling for socioeconomic economic variables, that students attending schools located in some regions that stance in an intermediate level of development tend to achieve higher scores than

they would achieved if their schools were located in Área Metropolitana de Lisboa. These evidences are in line with research outcomes of Pereira & Reis (2012) and Ferrer et al. (2010), taking into consideration Portuguese and Spain PISA surveys, respectively. Pereira & Reis (2012) advance that regions where school contributes to increase achievement-gaps benefit from better structural educational outcomes, higher development and higher inequality in teachers' experience. The authors consider, despite the incipient development of school choice in the Portuguese educational system, that the large number and the higher heterogeneity of schools in advantage regions, namely in Área Metropolitana de Lisboa, could explain the observed increase of inequalities in educational achievement and opportunities. Our findings are relevant since they can signal a complex puzzle of educational effects.

In the case of the regions that stance in an intermediate level of development and exhibits a higher level of educational equity and achievement, the results can express a proactive role of schools' influence correlated with a lower socioeconomic-achievement gap. In turn, the evidences that schools tend to exacerbate educational achievement-gap call on the family and socioeconomic factors behind schooling process, as wel as they pose several challenges when we infer family and socioeconomic variables impact.

An important focal point is the role of educational resources provide by family and public schools and their interconnectedness. Indeed, one can observe a trend to consider family and public educational resources as close substitutes. If this relationship appears to be consistent in more disadvantaged regions (see e.g. Tomula & Celik 2009); OECD 2005 and Woessmann, 2004), the question is more problematic in the context of developed regions. Some authors have been highlighting two relevant correlations. First, the family investment in education is positively correlated with the increase of household income. Second, there is a significant room to a positive correlation between family educational resources and public educational resources. According to some authors (Autor, 2014, Byun & Kim, 2010, Carneiro, 2008 and Reardon, 2012), the propensity to invest in education is not constant between income groups. It varies according to the level of family income. Thus, even though public investment benefits all social groups, it tends to reinforce the educational resources of the middle and high-income groups even more (Busemeyer, 2012; Chiu, 2010 and Chiu & Khoo, 2005).

Another important focal point is related with the aggregating impact of family background structures. We use student socioeconomic status and it is measured by the ESCS index. This index integrates parental education, occupation, income and wealth. It allows us to assess the agregated association of all socioeconomic dimensions. However, burgeoning literature on social segregation sustains that social background is not a homogeneous analytical category and the different dimensions can exert different impacts (see e.g. Hällsten & Thaning, 2018 and Erikson, 2016). For example, which of the four dimensions affect the propensity to invest in education more? Are the different dimensions weak or strongly associated? Indeed, the aggregation procedure may hide some barriers in capture the influence of the different dimensions. It further might affect the examination of the channels through which family variables exert influence on schooling process and outcomes.

The review of empirical research literature on regional student achievement-gaps does not reveal a clear relationship between the level of decentralization and the existence of significant regional-achievement gaps across regions. The significant regional heterogeneity observed appears to be a common feature. In fact, the question whether more centralised or decentralised educational governance models could nurture a more heterogeneous regional educational profile emerges as an open question.

The Portuguese education system is a centralised one (OECD, 2014), which has been under a reform process. Gradually, it has been increasing decision-making at sub-national levels, namely both in school governance model and municipalities responsibilities, from pre-primary to lower secondary education. Our findings indicate that socioeconomics status of students and schools exert significant influence on student achievement. The educational empirical literature highlights that the combination of family educational resources and public school resources can enhance the effect of schooling. The question is whether school per si can mitigate the power influence of socioeconomics status in order to harmonise students' learning conditions. These findings reinforce the results of Pereira & Reis, (2012) research and their

implication policy sugestions that educational policy in Portugal should be school-family-community oriented. The design of local educational policies is a key anchor to address the identified problems. The decentralisation of educational policies oriented to a more prominent role of schools and local governments can be a step forward to reshaping and develop suited local educational governance structures to address the effects of disadvantaged backgrounds.

6. CONCLUSION

This study represents an early attempt to understand and quantify the role of regions on educational performance by applying multilevel statistical models to PISA 2015 Portuguese data. The research aims at to contribute to disentangle the role of families, schools and characterisitics of the regions on student achievement in order to support the design of educational-related policies.

We found that variation explained due to region differences is smaller than 5%. The regional variance explained by individual or school composition concerning the socioeconomic status ranges (depending on the domain) from 20% to 24%. This impact relies mainly on variability across schools. After controlling for socioeconomic status, some regions present higher regional fixed effect than the Área Metropolitana de Lisboa, showing that their students tend to achieve higher scores than it would be expected given their socioeconomic status and the region where they live.

As the most important evidences, the findings suggest: (1) the importance of the predictive power of socioeconomic variables operating through family background and the school composition; and (2) the limited influence of educational-regional environment. However, relevant questions stem from the prominent influence of socioeconomic status of students and schools on student achievement. In particular, we debate, on the one hand, the importance of educational resources accumulation and whether the links between family and public resources are rule for assumptions of substitution effects or complementary effects. On the another hand, we advance some policy implications, namely the suggestion that in Portugal educational policies should be school-family-community oriented. In turn, we observed that the decentralisation of educational policies seems to be required to pursuit suited local educational governance structures to address the effects of disadvantaged backgounds.

Despite the extensive literature on the role of family background and school characteristics, in much less extent on the region background, the understanding of exact channels through which they exert their influence on educational outcomes remains an open issue. The consideration of regional dimension could provide insights how regional differences concerning economic, social and public policies dynamics affect educational resources accumulation. In addition, it expands the avenues to examine channels through which educational resources exert their impact on educational performance.

REFERENCES

Agasisti, T., & Cordero-Ferrera, J. (2013). Educational disparities across regions: A multilevel analysis for Italy and Spain. *Journal of Policy Modeling, 35(6),* 1079-1102.

Agasisti, T., & Vittadini, G. (2012). Regional Economic Disparities as Determinants of Students' Achievement in Italy. *Research in Applied Economics, 4*(2), 33-54.

Arellano, A., & Fullerton, T. (2005). Educational Attainment and Regional Economic Performance in Mexico. *International Advances in Economic Research*, *11*(2): 231-242.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the 'other 99 percent. *Science, 344* (6186), 843–51.

Berry, C., & Glaeser, E. (2005). The divergence of human capital levels across cities. *Papers in Regional Science, 84*, 407-444.

Brasington, D. (2002). Differences in the Production of Education Across Regions and Urban and Rural Areas. *Regional Studies, 36*(2), 137-145.

Bratti, M., Checchi, D., & Filippin, A. (2007). Territorial Differences in Italian Students' Mathematical Competencies: Evidence from PISA 2003. *IZA Discussion Paper No.* 2603, February. Retrieved December 12, 2018, from https://www.iza.org/publications/dp/2603.

Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy*, *22*(52), 781–861.

Buchmann, C. (2002). Measuring family background in international studies of education: Conceptual issues and methodological challenges. In A. Porter and A. Gamoran, (Eds), *Methodological Advances in Cross-National Surveys of Educational Achievement* (Chapter 6, pp. 150-197). Washington, DC: The National Academies Press.

Busemeyer, M. (2012). Inequality and the political economy of education: An analysis of individual preferences in OECD countries. *Journal of European Social Policy*, *22*(3), 219–240.

Byun, S. & Kim, K. (2010). Educational inequality in South Korea: The widening socioeconomic gap in student achievement. *Research in the Sociology of Education*, *17*, 155-182.

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

Carnoy, M., Garcia, E., & Khavenson, T., (2015). Bringing it back home: Why state comparisons are more useful than international comparisons for improving U.S. education policy. *EPI Briefing Paper nº 410*. Retrieved July 21, 2019, from https://www.epi.org/publication/bringing-it-back-home-why-state-comparisons-are-more-useful-than-international-comparisons-for-improving-u-s-education-policy/

Caro, D. (2015). Causal mediation in educational research: An illustration using international assessment data. *Journal of Research on Educational Effectiveness*, *8*(4): 577-597.

Checchi, D., & Peragine, V. (2010). Inequality of opportunity in Italy. *Journal of Economic Inequality*, *8*, 429-450.

Chiu, M. M. (2010). Effects of inequality, family and school on mathematics achievement: Country and student differences. *Social Forces*, *88*(4), 1645–1676.

Chiu, M. M., & Khoo, L. (2005). Effects of resources, inequality, and privilege bias on achievement: Country, school, and student level analyses. *American Educational Research Journal*. *42*(4), 575-603.

Donato, L., & Ferrer-Esteban, G. (2012). Desigualdades territoriales en España e Italia. Nuevas evidencias a partir de la evaluación PISA 2009. *Revista Española de Educación Comparada*, *19*, 105-138.

Erikson, R. (2016). Is it enough to be bright? Parental background, cognitive ability and educational attainment. *European Societies*, 18 (2): 117-135.

Ferrão, M. E., Barros, G., Bof, A. M., & Oliveira, A. S. (2018). Estudo longitudinal sobre eficácia educacional no Brasil: Comparação entre resultados contextualizados e valor acrescentado. *Dados - Revista de Ciências Sociais*, 61(4), 265–300.

Ferrão, M. E., & Couto, A. P. (2013). The use of a school value-added model for educational improvement: A case study from the Portuguese primary education system. *School Effectiveness and School Improvement*, 25(1), 174–190. https://doi.org/10.1080/09243453.2013.785436

Ferrer-Esteban, G. (2011). Beyond the traditional territorial divide in the Italian Education System. Effects of system management factors on performance in lower secondary school. Retrieved January 19, 2019, from https://www.researchgate.net/publication/286442322_Beyond_the_traditional_territorial_divide_in_the_Italian_Education_System_Effects_of_system_management_factors_on_performance_in_lower_secondary_school

Ferrer, F., Valiente, O., & Castel, J. (2010). Los resultados PISA-2006 desde la perspectiva de las desigualdades educativas: la comparación entre Comunidades Autónomas en España. *Revista Española de Pedagogía*, *245*, 23-48.

Haahr, J., Nielsen, T., Hansen, M. & Jakobsen, S., (2005). *Explaining Student Performance. Evidence from the international PISA, TIMSS and PIRLS surveys*. Danish Technological Institute. Retrieved May 8, 2017, from http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/35920726.pdf

Hällsten, M., & Thaning, M. (2018). Multiple dimensions of social background and horizontal educational attainment in Sweden. *Research in Social Stratification and Mobility*, *56*, 40-52.

Hanushek, E. (2016). What Matters for Student Achievement: Updating Coleman on the influence of families and schools. *Education Next*, *16* (2), Spring, 18-26.

Hanushek, E., & Woessmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across Countries. *Economic Journal*, *116*, C63-C76.

Heyneman, S., & Loxley, W. (1989). The effect of primary school quality on academic achievement across twenty-nine high and low income countries. *American Journal of Sociology, 88*, 1162-1194.

INE, (2018). Inquérito às Condições de Vida e Rendimento e 2018. Retrieved December 15, 2018, from https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=315156875&DESTAQUESmodo=2

OECD, (2018a). *Education at a Glance 2018: OECD Indicators*. Paris: OECD Publishing.

OECD, (2018b). *OECD Regions and Cities at a Glance 2018*. Paris: OECD Publishing.

OECD, (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education.* Paris: OECD Publishing.

OECD, (2014). *Education Policy Outlook: Portugal*. Paris: OECD Publishing. Retrieved May 15, 2018, from http://www.oecd.org/education/EDUCATION%20POLICY%20OUTLOOK_PORTUGAL_EN.pdf

OECD. (2008). *Measuring improvements in learning outcomes*. Paris, France.

OECD. (2005). *Learning for tomorrow's world first results from PISA 2003*, Paris: OECD Publishing. Retrieved April 13, 2017, from https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/34002216.pdf

Pereira, M.C., & Reis, H. (2012). What accounts for Portuguese regional differences in students' performance? Evidence OECD PISA". *Banco de Portugal - Boletim Económico*, *3*(Inverno), 59-83.

Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H., & Rasbash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society: Series B*, *60*(1), 23–40.

Rasbash, J., Steele, F., Browne, W.J. & Goldstein, H. (2014). *A User'sGuide to MLwiN - v2.31*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Reardon, S. F. (2012). The widening academic achievement gap between the rich and the poor. *Community Investments: Summer*, 24 (2), 19–39.

Rodríguez-Pose, A., & Tselios; V., (2009). Education and income inequality in the regions of the European Union. *Journal of Regional Science*, *49*(3), 411-437.

Shapiro, J. M. (2006). Smart Cities: Quality of life, productivity, and the growth effects of human capital. *Review of Economics and Statistics*, *88*, 324-335.

Sicilia, G., & Simancas, R. (2018). Equidad Educativa en España: Comparación Regional a partir de PISA 2015. *Fundación Europea Sociedad y Educación*. Retrieved November 9, 2018, from http://www.sociedadyeducacion.org/site/wp-content/uploads/Equidad-Educativa-en-Espana.pdf.

SMF, (2017). *On course for success? Student retention at university*. Retrived January 21, 2017, from http://www.smf.co.uk/wp-content/uploads/2017/07/UPP-final-report.pdf.

Tomul, E., & Celik, K., (2009). The relationship between the students' academics achievement and their socioeconomic level: cross regional comparison. *Procedia Social and Behavioral Sciences*, *1*(1), 1199–1204.

Woessmann, L. (2009). *International Evidence on School Tracking: A Review*. CESifo DICE Report 1, Retrieved January 9, 2017, from https://www.ifo.de/DocDL/dicereport109-rr1.pdf

Woessmann, L. (2007). *Fundamental Determinants of School Efficiency and Equity, German States as a Microcosm for OECD Countries*. Retrieved January 9, 2017, from https://sites.hks.harvard.edu/pepg/PDF/Papers/PEPG07-02_Woessmann.pdf

Woessmann, L. (2004). How equal are educational opportunities? Family background and student achievement in Europe and the United States. *CESifo Working Paper No. 1162*. Retrieved January 9, 2017, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=528209

APPENDIX

Table A1. Mean and Standard Error of the variables by region

Regions	SES average	S.E.	School SES estimate	S.E.
Alentejo Central	-0.14	0.07	-0.14	0.07
Alentejo Litoral	-0.46	0.10	-0.47	0.10
Algarve	-0.45	0.09	-0.46	0.08
Alto Alentejo	-0.30	0.05	-0.30	0.04
Alto Minho	-0.28	0.03	-0.28	0.03
Alto Tâmega	-1.08	0.36	-1.08	0.37
Área Metropolitana de Lisboa	-0.07	0.05	-0.07	0.05
Área Metropolitana do Porto	-0.40	0.09	-0.40	0.09
Ave	-0.92	0.10	-0.92	0.10
Baixo Alentejo	-0.38	0.27	-0.39	0.30
Beira Baixa	-0.38	0.16	-0.38	0.14
Beiras e Serra da Estrela	-0.35	0.30	-0.36	0.29
Cávado	-0.72	0.16	-0.72	0.16
Douro	-0.34	0.64	-0.34	0.64
Lezíria	-0.13	0.16	-0.13	0.16
Médio Tejo	-0.55	0.13	-0.55	0.12
Oeste	-0.75	0.09	-0.75	0.09
R. A. Madeira	-0.79	0.15	-0.80	0.16
R. A. Açores	-0.84	0.03	-0.83	0.00
Região de Aveiro	-0.44	0.07	-0.44	0.07
Região de Coimbra	-0.36	0.24	-0.36	0.24
Região de Leiria	-0.21	0.11	-0.21	0.11
Tâmega e Sousa	-1.08	0.09	-1.09	0.08
Terras de Trás-os-Montes	-0.50	0.16	-0.50	0.18
Viseu Dão Lafões	-0.15	0.29	-0.15	0.29

Acknowledgments
The third author was partially supported by the Project CEMAPRE/REM - UIDB/05069/2020 - financed by FCT/MCTES through national funds.

Disclaimer
The views expressed in this article are purely those of the authors and should not be regarded as the official position of the European Commission.

Funding
This work was partially supported by the Fundação para a Ciência e a Tecnologia [CEMAPRE-UID/MULTI/00491/2020].