Does family socio-economic background influence educational outcomes in Portugal? A country data evaluation on income distribution and inequality over time

Abstract

The socioeconomic background has been a significant factor in policy decisions relating to education throughout the past few decades. This research investigates how Portuguese students' academic performance is impacted by their parents' educational backgrounds. According to earlier studies, a parent's educational experience has a considerable impact on their child's academic success. The income makeup of the household and better support for the children are what led to these findings. This research contributes to the body of knowledge by capturing the interaction effect between educational level and time period using a dummy variable for higher education and a time dummy. The Regression results supported a similar conclusion to more recent theories, according to which high parental education levels, have less of an *impact on their children's academic performance than they did in the past, although low education levels still have a sizable detrimental impact.*

Master Thesis Policy Economics

Author: Débora de Jesus Nunes

Student Number: 613184

Supervisor: Schindler, DS

Second assessor: Dr. S.V. Kapoor

Date of the final version: 9 of February 9, 2023

Table of Contents

1. Introduction

Education levels have increased in many nations as a result of the unrestricted access to education for the population after realizing the impact that higher education levels have on society. In these nations there have been social advancement and economic prosperity, as well as less inequality between different racial and ethnic groups, and minorities. The concerns about education's quality have replaced questions about education's quantity as the dominant topic of public discourse. It is challenging to judge the quality of education in a school or a nation and to try to make comparisons, but this is done frequently through student achievement, resulting in rankings. Educational achievement depends, however, on multiple factors such as family background, access to resources, and the educational opportunities available.

The role that socioeconomic background, specifically household setting and parental history, plays in modeling individuals' academic routes and abilities is an important consideration when examining why certain students present better results than others in test scores. The ability of parental education to predict their children's academic success is well accepted. Every assessment of the relation between a parent's socioeconomic situation and their child is positive and significant (Hertz et al., 2007). However, more recent studies refute previous research by demonstrating that, for a variety of reasons, the parental background does not have the same significant impact on one outcome as it did in previous decades. For example, a recent study by O'Connell and Marks (2022), using data from Britain, found that psychological factors including personality traits and cognitive ability are far more effective at explaining educational attainment (i.e., school grades) than SES (socioeconomic status). Similar findings were found in a subsequent study by Björn Boman (2022), who used Swedish data and included variables related to SES, and cognitive and non-cognitive abilities of students in Grades 6 and 8. A sizable portion of the grade variance was explained by a variety of variables, with SES having the weakest link with educational attainment and cognitive and non-cognitive ability having the strongest.

These findings are in line with contemporary reality. For example, students, particularly those from low-income families, can use technology to understand a subject (cognitive ability), depending on their intrinsic motivation to learn, rather than relying on their parents' educational background to instruct them (non-cognitive ability). It is difficult to dispute that parents today can still "pay" for success by spending more time with their kids and encouraging them to work hard, even though current theories maintain that these effects are less pronounced than they formerly were.

Does the socioeconomic status of parents still have a substantial influence on youngsters' academic results? Is it still as crucial today as it was back then? Enhancing this knowledge of the underlying mechanisms causing economic outcomes to persist through generations is also important for policymakers. When hereditary abilities are the primary contributor to intergenerational relations, policies can only lessen the correlation by favoring those who are less capable. This will improve efficiency and have a positive impact on society. However, when the correlation is because of the external stimulus that higher-educated parents can provide, the government can intervene by recreating this environment for children from lower-educated parents (Cohen, 2016).

It is important from a human capital point of view that children of parents with lower levels of education can achieve their full potential, having a spillover effect that eventually reduces inequality (Black et al., 2005).

Motivated by the discrepancies between modern and old approaches regarding the effect that SES has on the academic results of students, this research develops a multivariable regression model to further evaluate these effects. By attempting to provide intuition regarding the relevance of the effect of SES on academic achievement, which can have implications for policymaking in the education sector, the main research question it seeks to address is to what extent SES influences educational achievement in Portugal, specifically in the years 2009, 2012, and 2018. Furthermore, this approach aims to analyze the impact and endurance of socioeconomic inequalities across generations in Portugal, the role that high parental education levels and other socioeconomic characteristics play in shaping students' abilities and provide intuition for future research.

The main model discussed in this paper explores how parental background can affect students' outcomes in three main domains: mathematics, reading, and science. The data was obtained through school questionnaires, organized by OECD (Organization for Economic Co-operation and Development), where students were asked to answer questions regarding their school and home environment.

The model consists of a multi-variable OLS regression where some assumptions are made. As a measure of school quality, a variable representing class time in minutes is used for each of the three subjects. The variable choice is based on the reasoning that schools that provide more teaching time have more resources (more teachers and staff) and consequently higher quality that can positively impact student outcomes.

With this model, our findings are in accordance with the most recent theories: the parental background had a statistically significant effect on students' outcomes in the last decade (2009, 2012). However, in more recent years, this is no longer observed in this sample. When accounting for the 2018 data set, it is possible to conclude that this effect is no longer significant for students' outcomes. Higher levels of parental education do not appear to be significant anymore. In contrast, lower levels of parental education still impact students' outcomes, probably due to other factors highly correlated with parental education levels. For example, parents with lower educational levels may have precarious occupations with irregular work hours that prevent them from giving their children the attention they need, negatively affecting the home environment.

The rest of the paper is structured as follows: Section 2 summarizes the relevant literature; Section 3 explains the Portuguese educational system and shows a descriptive analysis of government expenditure in the studied years. Section 4 explains PISA (Programme for International Student Assessment) from OECD, and Portugal's participation. Section 5 establishes the main model and descriptive analyses of the Portuguese population. Section 6 presents the subsequent results and discussion. Finally, Section 7 summarizes the main findings and presents this research's main conclusions.

2. Literature Review

The relationship between education and human capital has its origins in the 1960s. Two primary variations of the neoclassical model explain the connection between human capital and education. The first one encourages people to develop productivity-related abilities while still in school so that those who invest in education are more productive and will be rewarded for it in the future in the job market (Becker, 1964). The second views education as a trait that subtly relates to production. Higher education, for instance, often does not provide skills that can be immediately transferred to the labor market, but it does assist people in becoming more trainable in the workplace, which lowers the cost of future training (Barth, 1977).

Despite their differences, both theories support the idea that education can help one succeed in the job market in the long run. It is critical to remember that developing skills is a continuous process. At any given time, a person's accomplishments and learning capacity are the consequence of their prior investments and experiences, highlighting the value of parental involvement in children's education and growth as means of fostering intergenerational mobility (Carneiro, 2007). In this research, a linear regression model was created based on the Coleman report (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld and York, 1966) to assess the extent to which individuals' academic achievement depends on their family's socioeconomic and cultural status, as well as and the school environment. His findings regarding the importance of families in the educational success of teenagers (15 years of age) are similar to Coleman's research: while the family backgrounds of those students are the most crucial aspect of home and school environments, it has been observed that school resources have a minimal impact on educational outcomes. Socioeconomic background is primarily responsible for differences in test scores, general student performance, early school leaving, entrance to university or college, and total educational achievement between students or schools. (Marks, 2013).

Furthermore, according to Carneiro and Heckman (2003), ability gaps between children from various socioeconomic classes start to show up very early in life, frequently as early as 1 or 2 years old. Studies on child development (e.g., Olson, S., 2012) emphasize the importance of distinct life stages for the development of various capacities. By employing imaging technologies to investigate brain activity in reaction to different stimuli, it was possible to better understand how the children's environment and interaction with others can affect their physical, emotional, and behavioral development. These skill gaps are significant and frequently widen over time, indicating that complementary human capital investments occur at various points in time.

Additionally, Dickson et al. (2016) with British longitudinal study on parents and kids in England and Wales, discovered that an increase in parental education had a positive impact on their children's test scores from early childhood through the end of the compulsory schooling age (16 years old). According to their research, parental education plays a significant role in intergenerational outcomes. In their conclusion, they defend that characteristic that motivate the perusal of higher education may also impact parental skills or ability that influence the children's home environment and dedication towards school, which consequently can affect children's' ability to pursue higher education.

It seems obvious that if a person is better prepared to study at a young age, they will learn more effectively in the future. In this sense, variations in academic performance are likely to result in disparities in future labor market outcomes.

Some studies also consider the genetic influences that family origins may have. For instance, twin pairs were used by Behrman and Rosenzweig (2002) to control for the effects of "nurture," and they found no evidence that mother schooling raised children's educational achievement. Their findings might imply that factors other than genetics, such as family environment and upbringing, play a significant role. In accordance with that research, Plug (2004) examined the educational results of adoptees in Wisconsin and discovered that for mothers, hereditary skills and assortative mating are significant factors in the transmission of education between generations.

Regardless of parental previous human capital investments, future prospects also affect one's decision to attain more years of education or to join the labor market. According to the literature, individuals' circumstances and efforts, as well as their access to educational opportunities both play a role in one's future decisions (Golley and Tao Kong, 2018). Individuals' circumstances may refer to a person's gender, family background, wealth, or socioeconomic status, whereas individuals' effort refers to the motivation and assertiveness they exert to obtain the desired outcome. These circumstances, according to Davies, S., and Guppy, N. (1997), are well-known when discussing white and middle-class students, who are more likely to benefit from academic advantages because they have better pre-college educations, attain higher test scores, and have a variety of home advantages that support academic success.

Through the choices of their partners, parents can also have an impact on their children's academic performance. It is critical to keep in mind that fertility levels and patterns are influenced by marriage creation and upkeep, which has an impact on population growth rates and generational structures (Mare, 1980). Assortative mating is the process through which spouses pair up based on socioeconomic status and other social characteristics (Mare, 2000). According to Godoy et al. (2008), when couples of comparable traits join up, positive assortative mating takes place, whereas negative assortative mating takes place when they pair up with someone of a different characteristic (also known as a random match). An example of positive assortative mating is when both individuals have the same education level, for example, they are both high-educated, or low educated. This lack of randomization in marriage may eventually have an impact on the traits of offspring. The types of marriages that take place between individuals with different social traits determine the family background of their offspring, which influences the social makeup of the following generation (Mare,2000). The hypothesis that changes in assortative mating are causing long-term increases in educational disparity was examined using data from the US by the author. Although the link between husband and wife's educational performance has increased over the past 50 years, it was shown to have little impact on educational disparity.

As previously indicated, assortative mating can raise reproduction rates among couples from comparable socioeconomic backgrounds. Additionally, it can also improve children's access to resources and developmental chances. Since both parents are probably engaged in similar professions or have the same levels of education, it can also result in more financial stability for families. This phenomenon is important to consider when examining the effects of socioeconomic background and child education as one of the consequences is that children from lower socioeconomic backgrounds may perform worse academically than children from wealthier socioeconomic backgrounds. Children from lower socioeconomic levels could have a variety of difficulties that could impair their academic achievement (Carneiro, 2007). These include a lack of resources at home, a lack of growth chances, and restricted access to high-quality education, which increases the risk that they will pursue a lower educational path. Parents with advanced degrees are frequently better equipped to give their kids the tools and encouragement they need to thrive in school. This entails giving them access to a top-notch education, career possibilities, and other tools that can enable them to realize their full potential.

But are these socioeconomic disparities of the same relevance in weighing the decision of obtaining more years of education as in the previous years? Recent research has shown that socioeconomic background, particularly to higher levels of education, is becoming less and less important. Its impact on adults' occupations and earnings is weak when considering cognitive ability (Marks, 2013). Indeed, in most developed countries where society can easily access education, individuals with high abilities are earning higher incomes compared to people with lower abilities, even when they have the same level of education. (Gregorio, 2002). For example, individuals with certain types of abilities such as creativity, problem-solving, or communication skills may be able to leverage these abilities to get better jobs and higher salaries, regardless of their level of education. Additionally, the different levels of education offered in developed countries (vocational training and traditional academic courses) also play an important role in determining how much an individual can earn for the same level of education. Giving students the chance to learn in methods that best suit their specific needs and interests, by providing different sorts of educational abilities, inspires individuals to realize their full potential and motivates them to learn more successfully. Additionally, it enables students from all backgrounds to attend high-quality education, regardless of their financial condition or socioeconomic status. These declining effects of socioeconomic background and the importance of cognitive ability support many modernization theories (O'Connell and Marks (2022) and Björn Boman (2022)).

Regardless of the arguments that contend that a student's familial history has no bearing on their academic performance, there is an interest in estimating the link between parents' socioeconomic status—as determined by their income, education, and race— and their children's academic achievements. There are not many studies that analyze the relationship between children's outcomes and parents' background characteristics in Portugal, and more research is needed since diverse societies may produce different outcomes. Portugal does not share the same traits as the US and the UK; for instance, in the three years under study, low-education parents predominated the sample, whereas the US and the UK are distinguished by a higher proportion of highly educated families and educational accomplishments.

As a result, we would anticipate similar associations in the case of Portugal, but in the opposite direction, that low parental education levels reduce children's school performance.

3. Portuguese Education System: An Overview

In Portugal, the schools (teachers) and students are involved in the decision-making process for the individual to pursue a higher (or not) level of schooling. Education is free and compulsory until 18 years old when students complete their year 12 (complete high school level). However, only one of these two prerequisites is required. Education is controlled mainly by the State via the Ministry of Education, but municipalities can make small decisions mainly regarding extracurricular activities. There is public and private schooling at all levels, including higher education.

Education in Portugal officially starts at the primary level (legally compulsory), ensuring from an early age that each student is enrolled at a level suitable for their ability. This level of education generally lasts around 4 years. At the end of their primary school year, school examinations, as well as national exams, are conducted but the results are merely informative and do not dictate whether a student can pursue the next level of schooling, that is basic education. Basic education lasts for 5 years.

At the end of basic education, in year 9, teachers help students decide which field of expertise they wish to pursue in high school by performing numerical and language tests to assess the student's strengths. There are four main fields of specialty to choose from, and it is an important decision since it will influence the national examination at the end of the high school level. Unlike the previous levels of education, these exams will determine the domain that can be pursued at a higher education level.

Moreover, to help prevent early school dropouts, some options issue an equivalent level of schooling (more technical), usually available at 15 years old.

After choosing the field, students pursue the high-school level which lasts for 3 years. Students take national exams in the two specialized topics of their field at the conclusion of the 11th grade. At the end of the 12th grade, students have again national exams, where one usually is for the Portuguese language, and the other is regarding the main subject of the course. The nationwide online application process for higher education allows students to submit their top-choice universities, and entrance is determined by prioritizing students' academic performance.

Admission to higher education relies on two parts: the average grades obtained in all subjects attended during the 3 years of high school, and the grade obtained on the national exams required by the university, which are related to the degree the student is applying for. The application grade to the university is the result of the weighted average of the two averages, with the weights chosen by each university. It ranges from 0 to 20, and the higher it is, the higher the chances of being admitted to the university.

Additionally, higher education is split into two main subsystems: university and polytechnic, which can be either public or private. The university subsystem is meant to be theoretically found and research focused. The polytechnic subsystem is career-focused and aims to offer greater practical training.

3.1. Role of the Government: expenditures over time

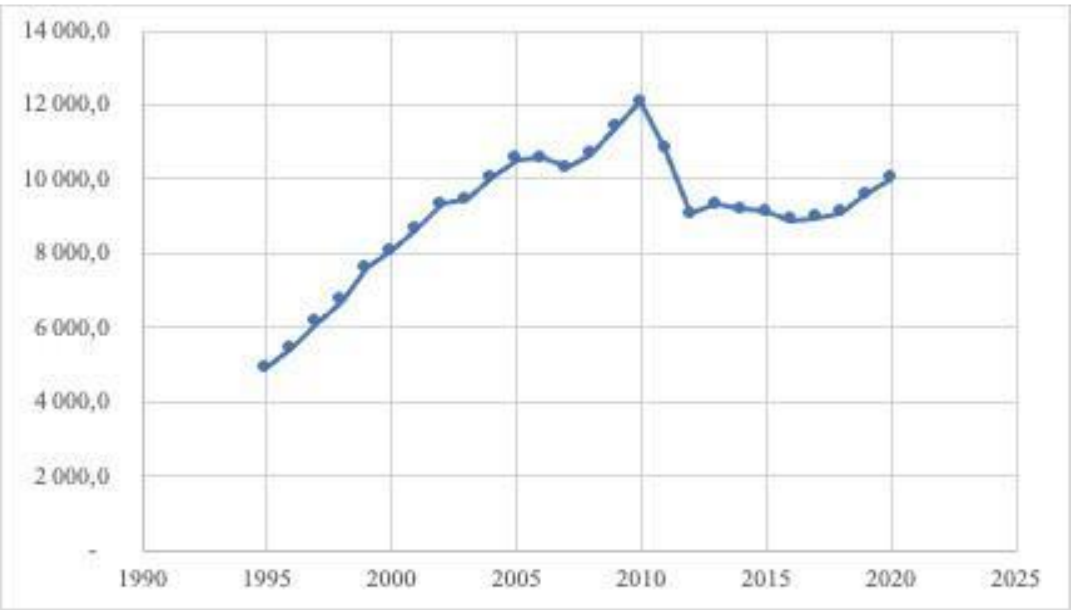
Economists, researchers, and policymakers have always been interested in education as an essential part of human capital. In this approach, it is crucial to take government spending into account when discussing educational success. The majority of the literature has demonstrated that investing in education has favorable externalities to society, both directly and indirectly, and that a significant portion of economic growth is attributable to the function of the accumulation of human capital (Köse and Güven, 2022).

In this sense, education increases labor force productivity, enhances overall welfare, and promotes growth as a major source of human capital (Jacobs and Bovenberg, 2009). The benefits of building up human capital and the disparity between social and private returns to education frequently serve as justifications for government intervention. The public sector typically provides the majority of funding for basic and secondary education, while student loans and scholarships are frequently used to subsidize postsecondary education (Dissou et al., 2016).

According to several studies, government spending on education increases growth lowers poverty and enhances overall well-being. However, it's important to keep in mind that this effect may not be instantaneous and that it can only show results in a period of 5 to 6 years when examining the effects of government investment on education.

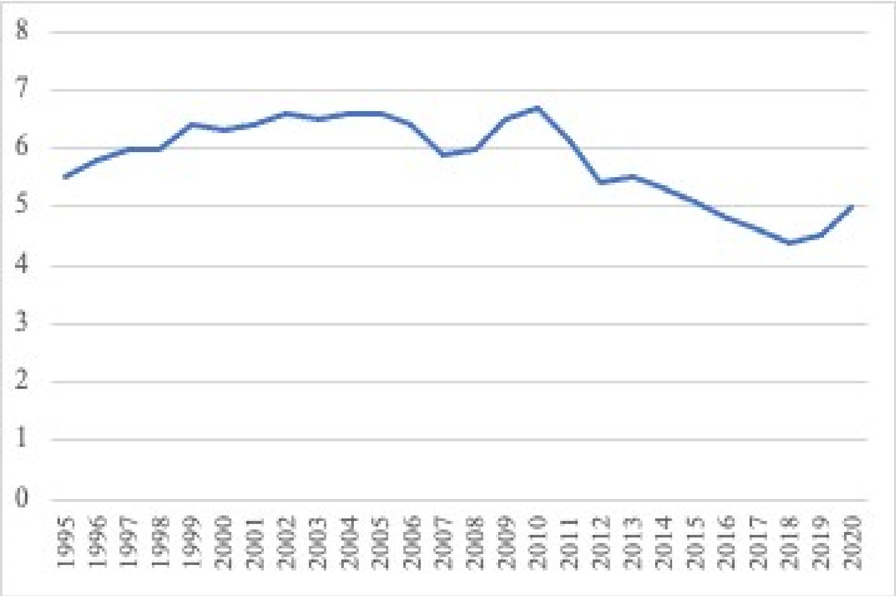
Considering this, the following graphs show Portugal's government expenditures on education in the last two decades.

Graph 1 - Public Administration Expenditure on Education by Portugal's government in millions between 1995 and 2020



Source: Despesas Das Administrações Públicas Em Educação, n.d., https://www.pordata.pt/portugal/despesas+das+administracoes+publicas+em+educacao-866, reproduced by author

Graph 2 - Public Administration Expenditure on Education as a % of GDP by Portugal's government between 1995 and 2020



Source: Despesas Das Admnistrações Públicas Em Educação Em % Do PIB, https://www.pordata.pt/portugal/despesas+das+administracoes+publicas+em+educacao+em+percentagem+do+pib-867, reproduced by the author

By looking at Graph 1, it is feasible to conclude that overall education spending increased during the period, rising from 4,927,50 million euros in 1995 to 10,015,70 in 2020. However, when comparing this expenditure to the country's GDP, it indicates a decline over the same time. Observing Graph 2, it is possible to divide it into the first and the second decade. In the first decade (1995-2005), it is possible to see a small growth going from 5,5% in 1995 to 6,6% in 2010. The highest increase in the ratio is visible between 2007 and 2010. when it increased from 5,9% in 2007 to 6,7% in 2010, reaching the maximum value of the two decades. This outcome can be ascribed to the 2008 financial crisis, which required additional social support from the government to help parents with their children's education.

In the second decade, it is possible to see a significant decrease in education expenditures in the country's GDP, decreasing from 6,7% in 2010 to 5% in 2020.

4. PISA

The data used in this thesis was obtained from the Programme for International Student Assessment, or PISA, which is organized by the OECD. PISA evaluates how well 15-year-olds can apply their reading, arithmetic, and science knowledge and skills to tackle real-world problems by solving problems.

PISA is one of the tools most frequently used to research educational policies (increasingly important in a growing number of countries). PISA has the most influence on how educational policies is determined and how the public perceives the quality of educational systems in other nations, both of which have been the focus of several studies (Breakspear, 2012; Carvalho, 2009; Figazzolo, 2009; Grek, 2009). In countries like France (Dobbins and Martens, 2012), Germany (Ertl, 2006), and Portugal (Afonso and Costa, 2009), among others, several authors have examined the effects of this OECD instrument (on the definition of educational policies.

The OECD conducts this survey every three years, encouraging students to complete a variety of questionnaires, including a background questionnaire in which they describe their traits, attitudes toward learning, and living arrangements (OECD, 2013). The data gathered by PISA assist nations and policymakers in examining relationships between student performance on PISA and background elements such as immigration, gender, socioeconomic status, and students' school and learning attitudes. To evaluate the degree of literacy in mathematics, reading, and science, cognitive tests are also conducted (OECD, 2013).

To better understand the meaning of literacy in each of the domains and according to the OECD (2016, p.15) defines each one as follows:

1. Scientific Literacy is the ability of an individual to engage in science-related issues and to understand scientific ideas, as a reflective citizen.
2. Reading Literacy is the ability of an individual to understand, use, reflect on and engage in reading written texts, to achieve their goals, to develop their knowledge and knowledge and their potential, and to participate in society.
3. Mathematical Literacy is the ability of an individual to formulate, apply and interpret mathematics in different contexts. It includes reasoning mathematically and using mathematical concepts, processes and facts, and tools to describe, explain and predict phenomena. It allows the individual to recognize the role of mathematics in the world and to formulate judgments and decisions based on reasons, as is expected of participatory, committed, and reflective citizens.

As a result, there are seven degrees of proficiency, with level 6 being the highest. It should be noted that PISA considers level 2 of proficiency to be the minimum level that all students should achieve. Below this level, PISA suggests, students lack the skills minimally required for active and effective participation in society.

Table 1 presents an overview of the six levels of science proficiency for scientific literacy along with variances between each level of proficiency. The score is estimated considering a scale from 0 to 1000 with a mean value of 500 and a standard deviation of 100. Additionally, PISA employs the imputation technique known as plausible values to estimate the values of a student's true ability based on their responses to the PISA assessment.

Please refer to appendix 1 for a more thorough explanation of what is required of the learner at each competence level.

Table 1 - Seven levels of proficiency in science in PISA 2015

	Scientific Literacy
Level 6	>=708
Level 5	>=633
Level 4	>=559
Level 3	>=484
Level 2	>=410
Level 1a	>=335
Level 1b	>=261

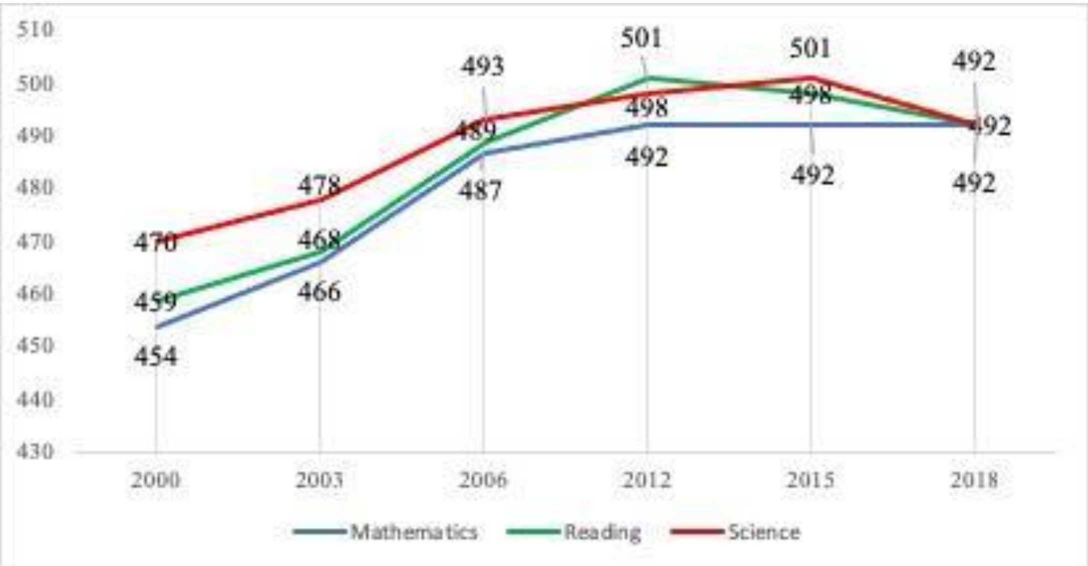
Source: PISA Test - PISA (n.d)

4.1. Portugal's participation in PISA

Portugal is one of the nations with a high percentage of participation in PISA from both schools and students, which eventually improves the sample's quality.

The average score among OECD countries is 500 points and the standard deviation is 100 points. Based on the collection of samples, it is possible to verify that Portugal has been progressing in the subject areas of science, reading, and mathematics.

Graph 3 - growth of Portugal in the different literacy domains evaluated by PISA between 2000 and 2018



Source: PISA results from 2018, https://www.oecd.org/pisa/publications/PISA2018_CN_PRT.pdf, produced by the author

From Graph 3, it is possible to observe, in all three courses, that Portuguese students have shown steady increases in their results over time. For instance, in mathematics, the improvement from 487 points in 2006 to 492 in 2012 and 2018 was roughly 5 points or a 0.05 standard deviation. Compared to mathematics, there was a greater improvement in reading. There was no improvement in 2018 from the scores of 489 and 498 in 2006 and 2012, respectively. The subject with the smallest improvement was science, which went from 493 in 2006 to 492 in 2018.

However, it should be noted that all three subjects had the highest improvement from 2003 to 2006. Science and Reading saw a modest fall between 2009 and 2012 (mathematics was the exception, stagnating), but there was a significant recovery in 2015. Overall, it is possible to see an upward trend of improvement in all three disciplines when measured over a longer time frame.

Table 2 shows that gender disparities in scientific literacy changed by year, from 2000 to 2015. When analyzing the results by gender, it is possible to confirm that Portuguese students, both male, and female, increased their outcomes in the three areas when results by gender are analyzed from 2000 to 2015. However, girls have been performing better in reading. The exact opposite has been true in mathematics, where boys have been outperforming girls.

Table 2 - Results by gender in mathematics, reading, and science between 2000 and 2015

		2000	**2003**	**2006**	**2009**	**2012**	**2015**
Mathematics	Female	446	460	459	481	481	487
	Male	464	472	474	493	493	497
Reading	Female	482	495	488	508	508	507
	Male	458	459	455	470	468	490
Science	Female	464	465	472	491	490	496
	Male	459	471	477	495	488	506

Source: PISA results from 2018, https://www.oecd.org/pisa/publications/PISA2018_CN_PRT.pdf, produced by the author

5. Empirical Analysis

5.1. Data

The observations used in this thesis refer to PISA data from the years 2009, 2012 and 2018. Each of those years includes data on students' and parents' characteristics in Portugal. The data structure is a pooled cross-section of PISA test results pertaining to students from the 7th through to the 11th grade and contains 5385, 5259, and 6023 observations for 2009, 2012, and 2018 respectively. It is important to note that the test was administered according to birth year rather than educational achievement, which accounts for the variation in educational attainment among the sample.

The majority of the factors used in this study are the ones that relate to the household's socioeconomic position, including those that regard the parents' backgrounds but also the student. The descriptive data for each year is displayed in Table 3 below.

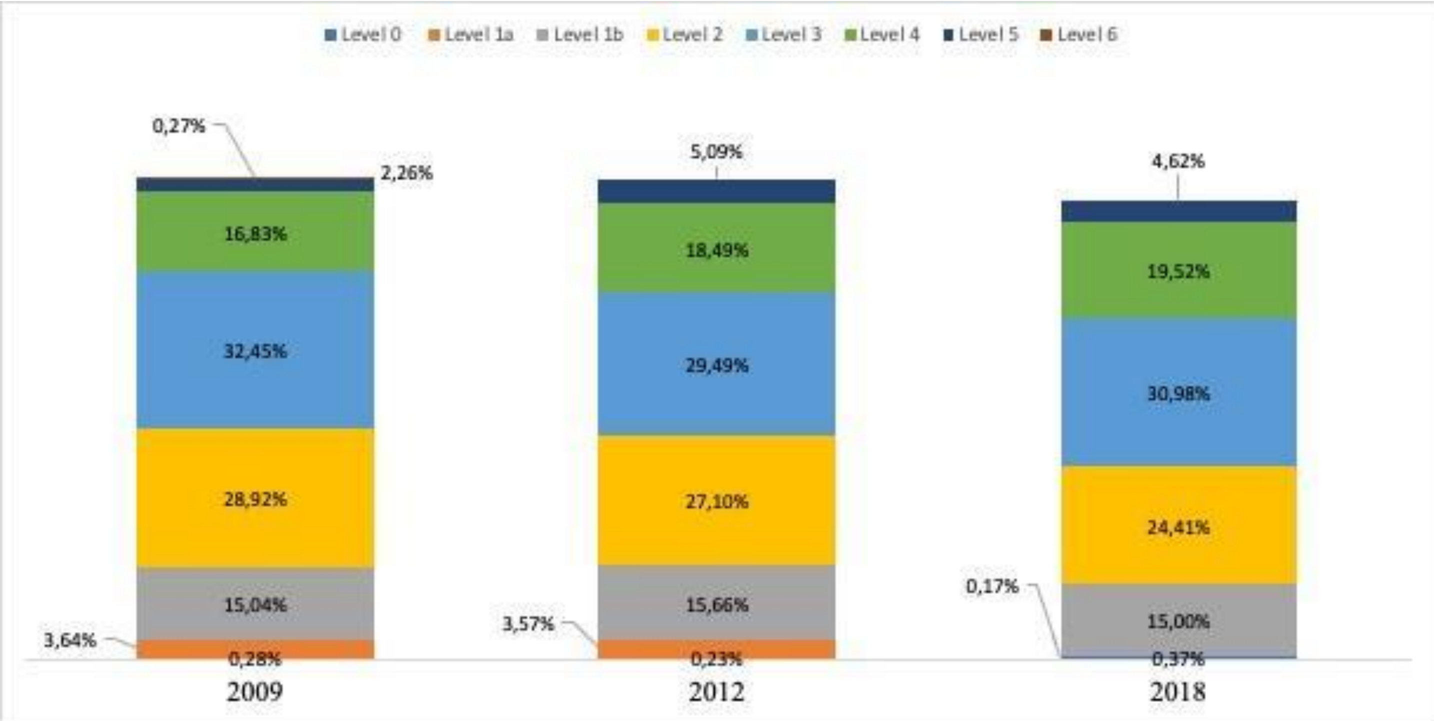
Table 3 - Descriptive Statistics for 2018, 2012, and 2009

	(1)	(2)	(3)
VARIABLES	2018	2012	2009
Math	496.495	493.150	489.252
	(2.802)	(3.503)	(2.970)
Read	495.239	493.699	491.330
	(2.577)	(3.553)	(3.224)
Science	494.757	495.337	495.228
	(2.931)	(3.484)	(2.918)
academic_year	9.577	9.515	9.493
	(0.019)	(0.028)	(0.031)
gender (yes=1)	0.498	0.496	0.513
	(0.008)	(0.008)	(0.006)
native (yes=1)	0.957	0.930	0.929
	(0.004)	(0.005)	(0.005)
repeat_student (yes=1)	0.253	0.320	0.339
	(0.012)	(0.018)	(0.019)
higher_educ_mother (yes=1)	0.399	0.226	0.212
	(0.012)	(0.014)	(0.011)

medium_educ_mother (yes=1)	0.211	0.222	0.208
	(0.007)	(0.007)	(0.007)
lower_educ_mother (yes=1)	0.390	0.551	0.580
	(0.011)	(0.017)	(0.013)
native_mother (yes=1)	0.852	0.868	0.873
	(0.007)	(0.008)	(0.006)
higher_educ_father (yes=1)	0.328	0.189	0.184
	(0.010)	(0.013)	(0.012)
medium_educ_father (yes=1)	0.200	0.187	0.194
	(0.006)	(0.006)	(0.007)
lower_educ_father (yes=1)	0.472	0.624	0.622
	(0.012)	(0.016)	(0.014)
native_father (yes=1)	0.849	0.884	0.891
	(0.007)	(0.007)	(0.006)
household_income_levels	2.046	-	3.139
	(0.046)	-	(0.065)
classtime_math_minutes	233.696	82.150	263.313
	(4.414)	(0.782)	(2.763)
classtime_read_minutes	211.490	82.239	224.885
	(3.681)	(0.804)	(2.475)
classtime_science_minutes	181.370	76.915	233.057
	(5.414)	(1.018)	(5.737)
Observations	5,385	5,269	6,023

The scores for each of the three subjects are represented by the variables math, read, and science. Graph 3 presents the distribution of the results in the sample of the three years to better assess where is the mean of the variables of the results.

Graph 4 - Percentage of students in each level of the OECD proficiency scale in mathematics for 2009, 2012, and 2018



From Graph 4 it is possible to observe that most of the sample is in levels 2,3, and 4 for mathematics, which represent proficiency levels between 410 and 559 score points. For more information regarding the proficiency levels please see annex A.1.

Regarding the variable household income levels, there is no data available for the year 2012. The information for 2009 and 2018 was collected from the parent's questionnaire, where parents were asked to self-select into one of six income levels, with level 1 being the lowest and 6 being the highest level.

Additionally, data privacy makes it impossible to know specific details about each student's school or teachers; therefore, the variable class time minutes for each course was added to account for school quality.

5.1.1. Parental background

The following figures, Tables 4, 5, and 6, will provide the proportion of students with both parents obtaining various degrees of education each year to help with understanding the level of education attained by Portuguese society. The interpretation of each table is as follows: in each column the proportion of students with mothers with low, medium, and high education levels, and each line the same type of information for the education level of the father.

Table 4 shows that in 2009 18.2% of students' fathers and 21.1% of students' mothers obtained higher-level education. Regarding medium-level education, it represents 20.3% of students' fathers and 22.3% of their mothers. Finally, low-level education parents dominated the sample with 61.5% and 56.6%, respectively.

Regarding Table 5, it is possible to notice a slight increase in the percentage of students with highly educated fathers and mothers, with 18.2% and 21.1% of the sample. However, the percentage of students' fathers with low education levels increased, representing 61.5% of the sub-sample.

In Table 6 more differences are observed, particularly a shift toward higher education, where it represents 32.4% of students' fathers and 40.6% of students' mothers. Medium-education fathers and mothers represent, respectively, 20.3% and 21.9%. Low-education parents still occupy a significant portion of the academic level with 47.3% of students' fathers and 37.6% of mothers.

By comparing the total proportion from 2009 to the upcoming years of the sample, it is possible to see that parental education in the last nine years has shifted positively towards a higher level of education. Nonetheless, the proportion of low-education parents is still the most significant in the 2018 sample.

Table 4 - The proportion of Students with Both Parents Attaining Various Levels of Education in 2009

	Mother's Education Level (% of students)			
Father's Education Level (% of students)	**Low - education**	**Medium- education**	**High- education**	**Total (%)**
Low education	.474	.104	.048	.615
Medium education	.051	.095	.043	.203
High education	.022	.032	.132	.182
Total (%)	.566	.223	.211	1

Table 5 - The proportion of Students with Both Parents Attaining Various Levels of Education in 2012

	Mother's Education Level (% of students)			
Father's Education Level (% of students)	**Low - education**	**Medium - education**	**High - education**	**Total (%)**
Low education	.477	.097	.041	.626
Medium education	.067	.095	.04	.189
High education	.021	.031	.130	.185
Total (%)	.547	.231	.222	1

Table 6 - The proportion of Students with Both Parents Attaining Various Levels of Education in 2018

	Mother's Education Level (% of students)			
Father's Education Level (% of students)	**Low-education**	**Medium-education**	**High - education**	**Total (%)**
Low education	.298	.1	.075	.473
Medium education	.046	.083	.074	.203
High education	.031	.035	.257	.324
Total (%)	.376	.219	.406	1

5.2. *Methodology*

An OLS regression will be carried out to assess the degree to which parental background has an impact on pupils' performance by examining the effects of these numerous variables on student accomplishment. Since PISA collects rich information on cognitive skills, family, and school environment, the data allows us to analyze the effect of these different variables on student achievement.

The first step will be to quantify how socioeconomic status affects test scores in the three subjects of math, science, and reading. The analysis starts with the investigation of the following model,

$$TS_{it} = \alpha + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 T_{it} + \beta_4 S_{it} + \beta_5 Z_{it} * T_t + \varepsilon_{it} \quad (1)$$

where TS_{ij} is a test score (in reading, mathematics, and science) for individual i in year t, X is a matrix representing student i's characteristics at time t, Z is the matrix with student i's parents' characteristics at time t, S is the matrix with student i's school characteristics at time t T_t as a time dummy, ε_{it} is the error term.

Matrix X of the students' characteristics includes academic_year, age, gender, native, repeat student (a binary variable that takes the value 1 if a student repeated a year, and 0 otherwise), and class time in minutes.
Matrix Z of parents' characteristics considers three binary variables that represent the level of education completed by each parent per ISCED levels (International Standard Classification

of Education) - lower_education, medium_education, and higher_education - two binary variables that assume the value 1 if each student's parent is a native Portuguese speaker, and 0 otherwise (native_mother and native_father), and household_income_levels. Matriz S will have the class time in minutes regarding each subject.

First, each year's regression is performed to check the consistency of the results. Second, since the main goal is to determine the significance of the effect of parental background over time, the interaction term in equation 1 between the higher education dummy variable and the time dummy will be used to capture the interaction effect of time on the dependent variables for students with parents with different educational levels. The equation will be replicated twice. First by using t=0 referring to 2012 as the baseline, and t=1 referring to 2009, and second by using t=0 as 2012 and t=1 referring to 2018.

Furthermore, the effects of paternal and maternal education are estimated separately because partners' education tends to be highly correlated. Higher-educated females typically marry men who are higher educated as well (Mare,1980). This dataset shows a correlation of 0.577 between paternal and maternal education. For this reason, equation 1 will be split into equation 1.a which refers to controlling for mother education, and 1.b. which refers when controlling for father education. Only levels of high education and low education will be considered to assess the disparities of the impacts at the most extreme level.

5.3. Hypotheses

Educational theory holds that households with high socioeconomic status (SES) are more able to give their children significant resources, not only financially (e.g., can afford additional tutoring) but also in terms of resources and opportunities available in the home environment (Holm et al., 2019).

Furthermore, children from high-SES homes will benefit from advantages, whether quantitative or qualitative, enabling them to take part in more superior university-level courses or any other learning opportunities that enhance their developmental outcomes (Raftery and Hout, 1993; Lucas, 2001). Given the available data and social stratification theories, it makes sense that children from high SES families would outperform in the three subjects under study. However, given the demographics in Portugal (the sample is dominated by parents with low levels of education), it is anticipated that the opposite effect will be confirmed. Children from low SES backgrounds would underperform in comparison.

Therefore, H.1 is our first hypothesis.

H.1: Children with parents from low-SES backgrounds have more pronounced differences in academic achievement.

Furthermore, individual personality features may have a larger effect on an individual's academic achievement than their family background, according to more recent education theories. Although it is difficult to consider such factors with this dataset, if we believe that students who fail a year are less driven, this variable (repeat_student=1) would have a large negative impact on the outcome variable than other variables related to parental background.

Therefore, H.2 is the second hypothesis

H.2: Students that repeated a year (repeat_student=1) are more demotivated and underperform. This variable impacts the outcome variable more significantly than low parental background.

Finally, continuing with the more recent educational theories, in more recent years (2018), highly educated parents should not have an impact in the outcome variables as they had in the past (years 2009 and 2012).

Lastly, H.3 is the last hypothesis

H.3: Parents with high levels of education do not significantly affect students' academic scores in 2018, although the opposite could be verifiable in 2009 and 2012.

6. Results

The results from the OLS regression regarding equation 1 will be shown in Tables 7, 8, and 9. This regression mainly studies whether parental education backgrounds affect students' performance each year and over time, for the different levels of education for the mother and father. The following conclusions will be for equation 1.a when we control for mothers' education level. Similar inference can be made for the father's education level. Please refer to the appendix for the tables regarding equation 1.b.

Table 7 – Results of the OLS regression for 2009 regarding equation 1.a

	(1)	(2)	(3)
VARIABLES	Math	Read	Science
academic_year	34.472***	31.575***	29.829***
	(3.434)	(3.746)	(3.258)
gender	-25.514***	12.981	-10.095***
	(2.988)	(23.190)	(3.024)
repeat_student	-48.727***	-46.892***	-41.877***
	(4.749)	(5.302)	(4.806)
native	-1.890	1.570	3.221
	(6.559)	(7.291)	(5.820)
household_income_levels	7.987***	7.674***	6.378***
	(1.191)	(1.063)	(1.057)
higher_educ_mother	18.477***	11.511*	18.746***
	(5.208)	(6.179)	(5.076)
lower_educ_mother	-6.689*	-6.621*	-6.331*
	(4.033)	(3.868)	(3.631)
native_mother	-0.134	-1.253	-0.265
	(5.075)	(4.330)	(4.368)
native_father	1.962	-0.573	0.902
	(5.623)	(5.445)	(4.475)
classtime_math_minutes	-0.027	-0.014	0.004
	(0.035)	(0.037)	(0.032)
classtime_read_minutes	0.035	0.023	-0.011
	(0.039)	(0.038)	(0.035)
classtime_science_minutes	0.060***	0.050***	0.047***
	(0.015)	(0.015)	(0.013)
Constant	169.180***	180.915***	214.980***
	(31.521)	(31.680)	(30.233)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8 – Results of the OLS regression for 2012 regarding equation 1.a

	(1)	(2)	(3)
VARIABLES	Math	Read	Science
academic_year	44.877***	39.798***	42.212***
	(3.323)	(3.774)	(3.654)
gender	-20.138***	28.730***	-5.577
	(3.117)	(3.804)	(3.502)
repeat_student	-60.133***	-44.648***	-46.543***
	(5.152)	(6.091)	(5.593)
native	8.457	5.697	3.681
	(7.374)	(7.450)	(7.596)
higher_educ_mother	16.796***	11.929***	11.455**
	(4.379)	(4.623)	(4.493)
lower_educ_mother	-15.319***	-17.023***	-15.641***
	(4.260)	(4.503)	(4.307)
native_mother	3.909	2.153	9.427*
	(6.610)	(6.024)	(5.016)
native_father	0.553	2.535	0.671
	(5.945)	(6.187)	(5.547)
classtime_math_minutes_	-0.651**	-0.309	-0.408
	(0.264)	(0.374)	(0.332)
classtime_read_minutes_	0.438*	0.299	0.285
	(0.254)	(0.369)	(0.303)
classtime_science_minutes_	0.090	0.010	0.039
	(0.103)	(0.118)	(0.120)
Constant	120.029***	130.155***	131.864***
	(31.942)	(35.440)	(35.395)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 9 – Results of the OLS regression for 2018 regarding equation 1.a

	(1)	(2)	(3)
VARIABLES	Math	Read	Science
academic_year	40.750***	40.639***	37.477***
	(4.683)	(4.037)	(3.821)
gender	-26.168***	8.329***	-20.378***
	(2.920)	(2.974)	(2.780)
repeat_student	-66.538***	-56.849***	-58.094***
	(8.968)	(7.066)	(6.429)
native	33.256***	28.254***	17.556**
	(8.544)	(6.951)	(7.740)
household_income_levels	5.089***	4.745***	4.584***
	(0.944)	(0.857)	(0.879)
higher_educ_mother	-1.412	-3.138	5.690
	(3.796)	(3.428)	(3.549)
lower_educ_mother	-10.837***	-8.944***	-3.949
	(3.356)	(3.426)	(3.769)
native_mother	0.718	-2.717	-1.312
	(4.134)	(4.250)	(4.570)
native_father	1.489	-2.609	1.093
	(3.777)	(3.989)	(3.552)
classtime_math_minutes	0.049***	0.002	0.017
	(0.016)	(0.013)	(0.014)
classtime_read_minutes	-0.108***	-0.051***	-0.082***
	(0.020)	(0.017)	(0.019)
classtime_science_minutes	0.060***	0.068***	0.073***
	(0.006)	(0.007)	(0.006)
Constant	106.345**	96.114**	143.577***
	(43.372)	(39.621)	(37.667)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.1. Results for OLS regression for 2009, 2012, and 2018

One of the critical conclusions from comparing the Tables for each year is that academic_year appears to have a beneficial impact on the outcomes in the three subjects (mathematics, reading, and science). In 2009, controlling for maternal education, an additional year of education would, on average, result in improvements in mathematics, reading, and science of 34.472, 31.575, and 29.829, respectively. With values of 44.877, 39.798, and 42.212, respectively, in 2012, the effect is slightly higher Additionally, a tendency towards improvement can be seen in 2018, with average effects of 40.750, 40.639, and 37.477, respectively. The variable is statistically significant at 1% in all three subjects in the three periods. This result supports the argument of Carneiro (2007), who states that learning is a cumulative process. According to this line of reasoning, education policy should consider a life-cycle perspective.

The outcomes of the regression also show that gender has a statistically significant impact on the results in the various subjects. Tables 7, 8, and 9 demonstrate that, on average, women outperform men in literature, but fall short in math and science. The three years under study indicate this pattern. In fact, in the subjects of mathematics and science subjects, females scored on average 25.514 and 10.095 points fewer than males in 2009. Similar to 2009, the figures show 20.138 and 5.577 points less, respectively. The results eventually got worse again in 2018, with women performing 26.168 and 20.378 points lower on average. In 2018, the variable is statistically significant at 1% in all three subjects, highlighting that mathematics is 1% statistically significant in the three periods. These science, reading, and math outcomes are in line with an OECD study for Portugal that was released in 2018. As stated in this research, "in all countries and economies that participated in PISA 2018, girls significantly outperformed boys in reading – by 30 score points on average across OECD countries. In Portugal too, girls scored on average 24 score points higher than boys. The gap was however lower than that observed in 2009 (38 score points)"(OECD, 2019, p.6).

Moreover, according to Tables 7, 8, and 9, the fact that the student is a repeat_student for at least a year significantly affects the results in all subjects. This pattern is observed in the three years of analysis: in 2009, repeat students scored, on average, 48.727, 46.892, and 41.877 points less than other students in mathematics, reading, and science, respectively. In 2012, the figures worsened for mathematics and science, to 60.133 and 44.648 points less, respectively, on average. Results in reading improved slightly to 44.648 points less, on average. In 2018, the size of this effect increases again, in all three categories, with repeat students performing 66.538, 56.849, and 58.094 points less, on average. The results of this variable's effect on the academic results in the three subjects illustrate the challenges faced by students who fail at least once. These challenges may be brought on, for instance, by a lack of ability or drive, which is hard to target at the individual level, but it may also be the result of teachers not succeeding to give their students the support and direction they need to fill their knowledge gaps. It is feasible to see that this variable has the most negative influence when compared to the others from the regression. In this sense, policy measures should try to focus on these students and try to provide a way to compensate them in case the results are driven because of the school's lack of support (for example, by providing extra tutoring).

Finally, the last control regarding student characteristics is native. The variable is statistically significant at 1% in the three subjects only in 2018. In 2018, a native student scored, on average, 33.256, 28.254, and 17.556 scores higher than other students in mathematics, reading, and science, respectively. The impact of this variable can easily be explained by native students having an advantage in learning and understanding the teachers and the materials taught in class. Additionally, although the variable that analyses the language spoken at home is not being considered in these models, this can also affect students' performance. According to Azzolini et al. (2012) immigrant students whose parents spoke their native language at home seemed to have a gap in their results.

For the two periods during which we have access to the data, the variable household_income_ levels is statistically significant at 1% for all three subjects. In 2009, if the student was one level of income higher would score, on average, 7.987, 7.674, and 6.378 scores higher than other students in mathematics, reading, and science, respectively. In 2018, the situation is similar for the three subjects, where students would obtain 5.089, 4.745, and 4.584 points higher, on average, respectively for each subject. It is important to point out that this variable has some limitations. In the questionnaires, parents were asked to select a level from 1 to 5 that better characterizes their household. This situation can be seen as a form of self-selection. Nonetheless, this variable is important from a policy point of view since most policy education measures target family income to reduce the gap between students from different income levels. Additionally, it is possible to see from an overall overview this variable is not the one that causes the biggest impact on student's results. As seen before, flunking a year has a higher negative impact in comparison with students belonging to a low-income family. This is important from a policy view since it can indicate that today's policies that are designed to target low-income families should be targeted on a school level and try to provide more overall support to the students, although there are no empirical foundations for this argument.

Regarding parental background, the three years follow the same trend: students with highly educated mothers (or fathers) tend to excel in all subjects, and students with low-educated parents tend to perform statistically significantly worse.

In 2009, students with highly educated mothers (higher_educ_mother=1) performed on average 18.477, 11.511, and 18.746 points higher in mathematics, reading, and science, respectively. This variable is statistically significant at 1% in mathematics and science, and 10% in reading. Students with highly educated fathers (high_educ_father=1) followed the same trend, performing, on average, 17.001, 15.449, and 15.673 points higher. This variable is statistically significant at 1% for all subjects.
Results associated with low-educated parents show the opposite effect. Students with low-educated mothers (lower_educ_mother=1) performed, on average, 6.689, 6.621, and 6.331 points less in mathematics, reading, and science, respectively, and students with low-educated fathers (lower_educ_father=1) performed, on average, 9.100, 5.204, and 11.331 points less.

Although the general trend continues in 2018 and 2012 as it did in 2009, the impact of high education is not that significant anymore. On the one hand, in 2018, students with high-education mothers performed, on average, -1.412, -3.138, and 5.690 points higher in mathematics, reading, and science, respectively. Students with high-educated fathers performed, on average, -0.569, 2.148, and 3.559 points higher. Neither of the variables is statistically significant at any level for any of the three subjects. On the other hand, low-educated parents continued to have a similarly negative effect on students' academic results.

Other variables that account for the parental background are if the parent is native or not. Surprisingly the variable does not seem to have a statistical effect on the results over the three years studied. This can be an indicator of positive reception from Portugal towards immigrants, by providing the necessary help to learn the language and integrate in society.

As previously indicated, this model tried to consider school quality by using the length of courses for the three subjects. The coefficient is low in all three disciplines, which is logical given that an additional minute of instruction cannot significantly impact a student's score, but it is noteworthy to note that, in 2009 the variable classtime_science is statistically significant at 1% in all subjects. In 2012, the variable classtime_mathematics is statistically significant at 5% for mathematics, classtime_reading is statistically significant at 10% for mathematics and classtime_science is not statistically significant at any level for any of the subjects. In 2018, the variable classtime_mathematics is statistically significant at 1% for mathematics, classtime_reading is statistically significant at 1% for all three subjects and classtime_science is statistically significant at 5% for mathematics and reading and at 1% for science.

The difference in the results over the three years can be justified by a general decrease in the student's performance and other outside factors not being considered in the regression. For example, the 2009 financial crisis resulted in lower education spending as shown in Graphs 1 and 2, although this event it is not controlled in the regression, although it can have a negative effect on education performance.

6.2 Results for Interaction term β_5

To better assess how parental background affects test scores over time, despite the individual analysis for each year, it is important to analyze the term β_5 from equation 1. The explanation for β_5 will be applied to equation 1.a, similarly to the previous results. A similar interpretation can be drawn regarding the fathers' educational background.

To calculate the overall effect of the interaction term it is necessary to compute the predicted value of the outcome variable for parents with and without higher education at different periods. From equation 1:

$$TS_{it} = \alpha + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 T_{it} + \beta_4 S_{it} + \beta_5 Z_{it} * T_t + \varepsilon_{it} \qquad (1)$$

It is possible to compute these as:

$$E(TS_{it}|Z_t = 0, T_t = t) = \alpha + \beta 1 + \beta 2 * 0 + \beta 3 t_t + \beta 4 + \beta 5 t_t * 0$$

$$= \alpha + \beta 1 + \beta 3 t_t + \beta 4$$

$$E(TS_{it}|Z_t = 1, T_t = t_t) = \alpha + \beta 1 + \beta 2 * 1 + \beta 3 t_t + \beta 4 + \beta 5 t_t * 1$$

$$= \alpha + \beta 1 + \beta 2 + \beta 3 t_t + \beta 4 + \beta 5 t_t$$

The difference between the predicted values for parents with and without higher education at different periods gives the overall effect of the interaction between the two dummies on the outcome variables.

Hence, the overall effect is

$$\begin{aligned} &\mathrm{E}(\mathrm{TS}_{\mathrm{it}}|\mathrm{Z}_{\mathrm{t}} = 1, \mathrm{T}_{\mathrm{t}} = \mathrm{t}_{\mathrm{t}}) - \mathrm{E}(\mathrm{TS}_{\mathrm{it}}|\mathrm{Z}_{\mathrm{t}} = 0, \mathrm{T}_{\mathrm{t}} = \mathrm{t}_{\mathrm{t}}) = \\ &\qquad = \alpha + \beta 1 + \beta 2 + \beta 3\mathrm{t}_{\mathrm{t}} + \beta 4 + \beta 5\mathrm{t}_{\mathrm{t}} - (\alpha + \beta 1 + \beta 3\mathrm{t}_{\mathrm{t}} + \beta 4) \\ &\qquad = \beta 2 + \beta 5\mathrm{t}_{\mathrm{t}} \end{aligned}$$

For instance, if the difference is positive, it would imply that those with higher education have a higher effect of time on the dependent variable. On the other hand, a negative interaction term would suggest that students with parents with higher education have a weaker influence of time on the dependent variable compared to parents without higher education.

The effects from equation 1 are presented in Table 10.

Table 10 – Results of the OLS regression with the interaction term regarding equation 1.a

	(1)	(2)	(3)
VARIABLES	Math	Read	Science
academic_year	41.873***	37.310***	37.871***
	(2.275)	(2.227)	(2.220)
gender	-24.233***	19.208	-8.765***
	(2.371)	(13.939)	(2.070)
repeat_student	-55.646***	-46.886***	-44.669***
	(3.550)	(4.936)	(3.865)
native	5.710	7.210*	5.855
	(4.411)	(4.339)	(3.912)
higher_educ_mother	18.154***	12.658***	12.443**
	(4.932)	(4.863)	(4.959)
lower_educ_mother	-16.814***	-16.585***	-17.068***
	(4.346)	(4.256)	(4.410)
native_mother	1.467	0.093	3.120
	(3.470)	(3.186)	(2.998)
native_father	5.354	4.275	3.222
	(3.618)	(4.343)	(3.168)

classtime_math_minutes	-0.026	-0.014	0.002
	(0.026)	(0.028)	(0.023)
classtime_read_minutes	0.036	0.021	-0.002
	(0.029)	(0.029)	(0.026)
classtime_science_minutes	0.044***	0.047***	0.032***
	(0.012)	(0.013)	(0.011)
Year (2009=1)	-15,844	-14.195***	-11.277**
	(5.643)	(5.384)	(4.553)
2009 * higher_educ_mother	7.653	6.287	12.226*
	(7.541)	(8.624)	(7.335)
2009 * lower_educ_mother	1.392	1.684	4.073
	(5.904)	(5.709)	(4.961)
Year (2018=1)	-18.101***	-16.796***	-24.791***
	(4.541)	(4.516)	(5.202)
2018 * higher_educ_mother	-12.581**	-9.997*	-0.179
	(6.396)	(6.048)	(6.486)
2018 * lower_educ_mother	2.782	1.247	7.958
	(5.102)	(5.277)	(5.842)
Constant	135.144***	152.972***	165.737***
	(21.446)	(22.617)	(21.621)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The overall effect of children with lower-educated mothers scored on average 15.422 points, 14.901 points, and 12.995 points lower on mathematics, reading, and science scores respectively, in 2009 compared to pupils overall in 2012. On the other hand, in 2009 compared to 2012, students whose mothers attained academic accomplishment scored on average 25.807 points better in mathematics, 18.945 points higher in reading, and 24.669 points higher in science.

For 2018, similar conclusions can be made. In comparison with students' performance in 2012, children with lower-educated moms overall scored 14.032 points, 15.338 points, and 9.110 points lower on math, reading, and science scores. Whereas, for higher levels of mother education, students scored on average 5.573 points better in mathematics, 2.661 points higher in reading, and 12.264 points higher in science.

Accordingly to these effects, students whose parents have low educational levels perform worse over time, and students whose parents have higher education do not outperform others significantly over time. When regressions are run individually for each period, people with higher educational backgrounds increase scores by a range of -3.138 to 18.746, and when we use the time interaction term the effect is smaller, indicating that the influence of parental background has diminished with time.

Although there are some limitations to the implications that can be drawn from these findings, they are consistent with more contemporary theories. As it was seen, socioeconomic position affects the three subjects' literacy scores. However, individual students' characteristics present a larger effect (for example, gender, repeater status, and age have a higher influence on literacy scores than the socioeconomic background in this data set). A more general interpretation of the findings would imply that personal traits, such as ability, motivation, maturity, and even learning techniques (although these are not examined in this model), can significantly impact a student's academic success. By better understanding these personal characteristics of students, teachers can better adapt their teaching strategies to guarantee that every student can achieve their full potential.

7. Conclusion

There has been a lot of research performed on how parents' education affects their children's schooling. Although it is widely agreed that parental education benefits children's schooling, it is less clear what factors cause this relationship. To evaluate the causal effect of parental education, various methodologies are applied. This thesis contributes to the existing literature by attempting to estimate the causal effect of parental education in Portugal.

The study conducted for this dissertation had two primary goals: (1) to compare student literacy performance over the years by identifying and investigating its key drivers, (2) to assess if those drivers remain to have an impact in more recent years. The dependent variable in this work is the scores of each student in Mathematics, Reading, and Science literacy.

This research exploits the effect of the PISA results from three years on parental and student backgrounds. This effect is exploited by using a multi-variable OLS model to achieve the intended objectives. In this paper, we showed that parents' high academic achievement does not guarantee success, although poor parental background still has a statistically significant negative effect in the three presented subjects. These results are related to H.1.

For example, in 2009, a student with a highly educated mother would perform on average 18.477, 11.511, and 18.746 points higher in mathematics, reading, and science, respectively. Whereas, in comparison, the students with lower educated mothers performed, on average, 6.689, 6.621, and 6.331 points less in the same subjects, respectively. Moving forward to 2018, we see the same trend overall, where students with highly educated parents outperform students with lower-educated, although this variable is not statistically significant anymore. These findings also appear to support H.3, which predicted that the effect of high parental socioeconomic background on childrens' academic results would not be substantial.

Additionally, this Hypothesis seem to be aligned with the new theories where socioeconomic status does not have the same significant impact on students' academic outcomes. From a policy point of view, this is an important result since it shows that nowadays more population is getting instructed and that the students' school outcomes are not impacted anymore as they were in previous generations.

Despite the disagreements between modern and old theories, it is crucial to note that socioeconomic status can still have a negative impact on inequality and education, even if that impact is lessening, especially when referring to lower socioeconomic status. The importance of certain financial considerations in influencing a person's decision to seek higher education is reinforced by several perspectives. For instance, parents' increasing ability to pay for "success" (private schools, more tutoring) may have an impact on students' capacity to earn better grades, which may have an impact on students' decisions to pursue higher education levels (Gregorio, 2002). Additionally, an individual's inherent ability is one of the factors that cause more discussion regarding educational achievement, whether it is due to genetic or environmental factors that cause some students to outperform others in school. Another important factor, that is also hard to account for, is the peer effect. Peer effects can be summed as student A educational outcomes being affected by having student B as either a classmate or schoolmate, by student B's academic performance motivating the teacher, for example, or by student B's academic performance itself. If we consider that wealthy families can afford private, instead of public, education this influence may manifest and influence one's surroundings and networking. Students from these two types of education can be exposed to very distinct environments, where teachers' performance can also be significantly different (due to different monetary incentives), which can impact their academic performance.

Furthermore, cultural aspects might favor people from high socioeconomic backgrounds or put obstacles in the way of people from poor socioeconomic backgrounds.

Despite the consequences of socioeconomic background, it is important to consider different policy approaches that focus on schools and school quality rather than the socioeconomic background of the students. An example of a policy with emphasis on educational institutions and educational quality rather than socioeconomic background is a school choice or school voucher program.

Families are given public funding through a school voucher scheme so they can pay for their preferred private school's tuition. These programs provide parents with more options and more influence over where their children attend school to create competition between schools and raise educational standards. School voucher programs attempt to raise the standard of education regardless of the socioeconomic status of the family by giving parents the ability to select a private school. The argument is that schools will enhance their performance to attract more families by competing for students.

Another illustration of a strategy that concentrates on schools and educational quality rather than social background is a school reform program, such as the No Child Left Behind Act or the Race to the Top initiative. The goal is to help students perform better and hold teachers more accountable, by giving funds and support to schools. Regardless of the socioeconomic status of the pupils, these programs strive to raise the quality of education by establishing performance criteria for students and holding schools responsible for achieving those requirements. This can involve actions like giving failing schools more funding, developing novel teaching techniques, or providing professional development for the professionals.

7.1 Data Limitations

Multilevel regression models are used to analyze the links between school and system factors and student accomplishment. While these models are suitable for estimating connections in a way that considers the characteristics of PISA data, they lack the evidence to make causal conclusions, which most likely causes the results to be overestimated. As was previously mentioned, knowledge and skill are developed throughout a lifetime. Some factors that can have an impact on the outcomes are not considered in PISA statistics. For instance, as the National Research Council et al. (2002) noted, it is crucial to consider the context of the learning opportunity because it can affect achievement, for example, the school environment.

While it was attempted in this study to include some controls in the regression models, most notably the socioeconomic status and academic performance of the students, these controls alone are unlikely to account for all potentially confounding variables. The possibility of unmeasured factors, affects, for instance, the understanding of relations at the level of the school, the country, and the individual student. A peer effect, for instance, as mentioned before, may indicate that a student is (on average) better off in a private school, even if some of the benefits come from being around wealthy peers rather than from better instruction. However, PISA data does not take this into account (Hamilton and Corporation, 2009).

8. Reference list

Adani, N. (2017, December 13). The Effect of Parental Education on Students' Educational Outcomes over Time in the Netherlands. Business Economics.

Afonso, Natércio, Estela Costa (2009). The influence of the Programme for International Student Assessment (PISA) on policy decision in Portugal: the education policies of the 17th Portuguese Constitutional Government. Sísifo. Educational Sciences Journal, 10, pp .53-64.

Azzolini, D., Schnell, P., and Palmer, J. R. B. (2012). Educational Achievement Gaps between Immigrant and Native Students in Two "New" Immigration Countries: Italy and Spain in Comparison. The ANNALS of the American Academy of Political and Social Science, 643(1), 46–77.

Barth, M. C., and Thurow, L. C. (1977). Generating Inequality: A Review Article. The Journal of Human Resources, 12(1), 92.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2004). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. SSRN Electronic Journal.

Behrman JR, Rosenzweig MR (2002), Does increasing women's schooling raise the schooling of the next generation? American Economic Review 92(1): 323–334.

Boman, B. (2022). The influence of SES, cognitive, and non-cognitive abilities on grades: cross-sectional and longitudinal evidence from two Swedish cohorts. European Journal of Psychology of Education.

Breakspear, S. (2012). The policy impact of PISA: an exploration of the normative effects of international benchmarking in school system performance. OECD Education Working Papers, 71, OECD Publishing.

Carneiro, P. (2007). Equality of opportunity and educational achievement in Portugal. Portuguese Economic Journal, 7(1), 17–41.

Carneiro, P. M., and Heckman, J. J. (2003). Human Capital Policy. SSRN Electronic Journal.

Carvalho, L. M. (2009). Governando a educação pelo espelho do perito: uma análise do pisa como instrumento de regulação. Educação &Amp; Sociedade, 30(109), 1009–1036.

Cohen, Christian. (2016, July 28). The Intergenerational Effect of Parental Education in Germany. Business Economics.

Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld and R. York (1966), Equality of Educational Opportunity, Washington, D.C.; U.S. Government Printing Office.

Davies, S., and Guppy, N. (1997). Fields of Study, College Selectivity, and Student Inequalities in Higher Education. Social Forces, 75(4), 1417–1438.

Dickson, M., Gregg, P., and Robinson, H. (2016). Early, late or never? When does parental education impact child outcomes? The Economic Journal, 126(596). s

Dissou, Y., Didic, S., and Yakautsava, T. (2016). Government spending on education, human capital accumulation, and growth. Economic Modelling, 58, 9–21.

Dobbins, Michael, e Kerstin Martens (2012). Towards an education approach à la finlandaise French education policy after PISA. Journal of Education Policy, 27 (1), pp. 23-43.

Ertl, H. (2006). Educational standards and the changing discourse on education: the reception and consequences of the PISA study in Germany. Oxford Review of Education, 32 (5), pp. 619-634.

Figazzolo, L. (2009). Impact of PISA 2006 on the education policy debate. Education Internat.

Gary S. Becker. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. Social Science Research Network.

Godoy, R., Eisenberg, D. T., Reyes-García, V., Huanca, T., Leonard, W. R., McDade, T. W., and Tanner, S. (2008). Assortative mating and offspring well-being: theory and empirical findings from a native Amazonian society in Bolivia. Evolution and Human Behavior, 29(3), 201–210.

Golley, J., and Kong, S. T. (2018). Inequality of opportunity in China's educational outcomes. China Economic Review, 51, 116–128.

Gregorio, J. D., and Lee, J. W. (2002). Education and Income Inequality: New Evidence From Cross-Country Data Review of Income and Wealth, 48(3), 395–416.

Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. Journal of Education Policy, 24, pp. 23–37.

Hamilton, Laura S, and RAND Corporation. (2009). Using PISA Data to Measure School and System Characteristics and Their Relationships with Student Performance in Science. Unpublished.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2008). The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. The B.E. Journal of Economic Analysis &Amp; Policy, 7(2).

Holm, A., Hjorth-Trolle, A. and Jæger, M.M. (2019) Signals, educational decision-making, and inequality, European Sociological Review, 35(4), 447–460.

Jacobs, B., and Bovenberg, A. L. (2009). Human capital and optimal positive taxation of capital income. International Tax and Public Finance, 17(5), 451–478.

Köse, S., and Güven, A. (2007). Government education expenditures and income inequality: evidence from provinces of Turkey. SEER: Journal for Labour and Social Affairs in Eastern Europe, 10(1), 79-101.

Lucas, S.R. (2001) Effectively maintained inequality: Education transitions, track mobility, and social background effects, American Journal of Sociology, 106, 1642–1690.

Mare, Robert. D. (1980). Social Background and School Continuation Decisions. Journal of the American Statistical Association, 75(370), 295–305.

Mare, Robert D.(2000). Assortative Mating, Intergenerational Mobility, and Educational Inequality. California Center for Population Research. UCLA: California Center for Population Research

Marks, G. N. (2013). Education, Social Background and Cognitive Ability: The Decline of the Social. Routledge.

Marôco, J., Lourenço, V., Mendes, R., and Gonçalves, C. (2016). TIMSS 2015 – Portugal.Volume I: Desempenhos em Matemática e em Ciências. Lisboa: Instituto de Avaliação Educativa.

National Research Council, Education, D. O. B. A. S. S. A., Education, C. F., Assessment, B. O. T. A., Board on International Comparative Studies in Education, Gamoran, A., and Porter, A. C. (2002b). Methodological Advances in Cross-National Surveys of Educational Achievement (Illustrated). National Academies Press.

O'Connell, M., and Marks, G. N. (2022b). Cognitive ability and conscientiousness are more important than SES for educational attainment: An analysis of the UK Millennium Cohort Study. Personality and Individual Differences, 188, 111471.

OECD. (n.d.). PISA database instructions [Data]. OECD. https://www.oecd.org/pisa/data/httpoecdorgpisadatabase-instructions.htm

OECD (2012), PISA 2009 Technical Report, PISA, OECD Publishing, Paris.

OECD, (2013). PISA 2012 Assessment and Analytical Framework. OECD Publishing.

OECD, (2016). PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy, OECD Publishing, Paris.

OECD, (2019). PISA 2018 Results (Volume I): What Students Know and Can Do. OECD Publishing, Paris.

OECD. (2015, December 3). Summary description of seven levels of proficiency in science (PISA 2015) [Data]. OECD. https://www.oecd.org/pisa/test/summary-description-seven-levels-of-proficiency-science-pisa-2015.htm

Olson, S. (2012). *From Neurons to Neighborhoods: An Update--Workshop Summary*. National Academies Press. 500 Fifth Street NW, Washington, DC 20001.

Plug, E. (2004). Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. The American economic review, 94(1), 358-368.

pordata. (n.d.). Despesas das administrações públicas em educação [Data]. pordata. https://www.pordata.pt/portugal/despesas+das+administracoes+publicas+em+educacao-866

pordata. (n.d.). Despesas das administrações públicas em educação em % do PIB [Data]. pordata.https://www.pordata.pt/portugal/despesas+das+administracoes+publicas+em+educacao+em++percentagem+do+pib-867

Raftery, A. E. and Hout, M. (1993) Maximally maintained inequality: Expansion, reform and opportunity in Irish education, 1921–1975, Sociology of Education, 66, 41–62.

A. Appendix

A.1 Summary description of the seven levels of proficiency in science in PISA 2015

According to PISA 2015 (*PISA Test - PISA*, n.d.), the levels of proficiency are described below.

At Level 6 Students can use content, procedural, and epistemic knowledge to offer explanatory hypotheses for novel scientific phenomena, events, and processes or to make predictions. They can draw on a variety of interrelated scientific ideas and concepts from the physical, life, earth, and space sciences. They can distinguish between relevant and irrelevant information when assessing data and evidence, and they can draw on knowledge outside the scope of the typical school curriculum. They are able to discriminate between arguments supported by theories and evidence from those supported by other factors. Students at the sixth-grade level can compare several designs for difficult experiments, field investigations, or simulations and defend their decisions.

At Level 5, students are able to describe unfamiliar and more complicated facts, events, and processes with many causal links using abstract scientific ideas or concepts. They are able to employ theoretical knowledge to analyze data or make predictions, as well as more advanced epistemic knowledge to evaluate various experimental designs and support their decisions. Students at this level are able to assess several scientific approaches to a given subject and recognize the constraints placed on how data sets, including their origins and sources of uncertainty, can be interpreted.

At Level 4, students can create explanations of more complicated or abstract events and processes using content knowledge that is either presented or recalled. They are able to carry out experiments with two or more independent variables in a limited setting. With the use of procedural and epistemic knowledge, they can defend an experimental strategy. Students at this level are able to evaluate data from a somewhat complicated data set or an unfamiliar environment, come to suitable judgments based on those conclusions, and defend their decisions.

At Level 3, students can recognize or create explanations for well-known events using knowledge of somewhat complicated topics. They can create explanations with the proper cueing or support in situations that are more challenging or unfamiliar. To conduct a straightforward experiment in a confined environment, they can make use of procedural or epistemic knowledge components. Students at the third-grade level can differentiate between scientific and non-scientific concerns and recognize the supporting data for a scientific conclusion.

At Level 2, students can recognize a suitable scientific explanation, analyze data, and recognize the issue being addressed in a straightforward experimental design by using their everyday subject knowledge and fundamental procedural knowledge. They can get a reliable conclusion from a small body of evidence by applying fundamental or common scientific knowledge. Students at Level 2 can recognize topics that can be researched scientifically, demonstrating fundamental epistemic understanding.

At Level 1a, students can recognize or find explanations for straightforward scientific phenomena using basic or everyday content and procedural knowledge. They can conduct structured scientific investigations with no more than two variables with assistance. They have the ability to recognize straightforward causal or correlative links and to evaluate graphical and visual data with little cognitive load. Students at Level 1a may decide which scientific theory best explains presented data in familiar personal, local, and global situations.

Students at Level 1b may identify components of well-known or straightforward phenomena using fundamental or applied scientific knowledge. They can recognize elementary scientific phrases, simple patterns in data, and clear directions to perform a scientific technique.

A.2 Results for Tables 7,8,9 and 10 for equation 1.b.

Table 7 – OLS regression for 2009 regarding equation 1.b

	(1)	(2)	(3)
VARIABLES	Math	Read	Science
academic_year	34.822***	31.980***	30.088***
	(3.410)	(3.661)	(3.214)
gender	-25.582***	13.024	-10.210***
	(2.976)	(23.258)	(3.004)
repeat_student	-48.227***	-46.354***	-41.284***
	(4.737)	(5.301)	(4.805)
native	-0.503	2.915	4.525
	(6.579)	(7.263)	(5.922)
household_income_levels	8.137***	7.707***	6.388***
	(1.191)	(1.103)	(1.068)
higher_educ_father	17.001***	15.449***	15.673***
	(5.909)	(5.860)	(5.512)
lower_educ_father	-9.100**	-5.204	-11.331***
	(4.115)	(4.915)	(4.240)
native_mother	-1.465	-2.387	-1.324
	(5.068)	(4.256)	(4.236)
native_father	1.989	-0.384	1.146
	(5.496)	(5.408)	(4.479)
classtime_math_minutes	-0.031	-0.016	-0.001
	(0.036)	(0.037)	(0.032)
classtime_read_minutes	0.038	0.024	-0.007
	(0.039)	(0.039)	(0.035)
classtime_science_minutes	0.061***	0.051***	0.049***
	(0.015)	(0.015)	(0.013)
Constant	167.986***	175.679***	216.537***
	(31.315)	(30.535)	(29.795)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8 – OLS regression for 2012 regarding equation 1.b

VARIABLES	(1) Math	(2) Read	(3) Science
academic_year	45.472***	40.685***	42.731***
	(3.290)	(3.737)	(3.637)
gender	-19.864***	28.851***	-5.135
	(2.959)	(3.669)	(3.284)
repeat_student	-59.615***	-44.614***	-45.029***
	(5.254)	(6.037)	(5.556)
native	9.804	7.074	4.464
	(7.356)	(7.435)	(7.641)
higher_educ_father	13.385**	10.304	10.936*
	(6.240)	(6.293)	(5.917)
lower_educ_father	-19.860***	-16.691***	-21.965***
	(5.467)	(5.709)	(5.775)
native_mother	1.603	-0.310	7.776
	(6.545)	(6.108)	(5.090)
native_father	3.898	5.001	4.368
	(6.107)	(6.414)	(5.582)
classtime_math_minutes_	-0.663***	-0.328	-0.410
	(0.247)	(0.366)	(0.323)
classtime_read_minutes_	0.460**	0.326	0.299
	(0.234)	(0.356)	(0.289)
classtime_science_minutes_	0.093	0.014	0.037
	(0.102)	(0.119)	(0.120)
Constant	116.128***	121.373***	128.454***
	(31.635)	(35.457)	(34.498)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 9 – OLS regression for 2018 regarding equation 1.b

	(1)	(2)	(3)
VARIABLES	Math	Read	Science
academic_year	41.321***	41.070***	37.672***
	(4.757)	(4.058)	(3.795)
gender	-26.200***	8.371***	-20.381***
	(2.908)	(2.962)	(2.761)
repeat_student	-66.994***	-56.765***	-57.878***
	(8.965)	(7.064)	(6.447)
native	33.544***	28.628***	17.623**
	(8.682)	(7.002)	(7.826)
household_income_levels	5.319***	4.448***	4.571***
	(0.987)	(0.877)	(0.881)
higher_educ_father	-0.569	2.148	3.559
	(4.452)	(3.476)	(3.887)
lower_educ_father	-6.421	-5.891*	-7.468*
	(4.098)	(3.566)	(4.461)
native_mother	0.025	-3.014	-1.912
	(4.154)	(4.273)	(4.628)
native_father	2.029	-1.556	2.310
	(3.748)	(4.063)	(3.576)
classtime_math_minutes	0.049***	0.002	0.016
	(0.015)	(0.013)	(0.014)
classtime_read_minutes	-0.109***	-0.052***	-0.082***
	(0.019)	(0.017)	(0.019)
classtime_science_minutes	0.060***	0.068***	0.073***
	(0.006)	(0.007)	(0.006)
Constant	99.078**	89.166**	144.259***
	(44.577)	(39.848)	(37.037)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 10 – Results of the OLS regression with the interaction term regarding equation 1.b

VARIABLES	-1 Math	-2 Read	-3 Science
academic_year	42.054***	37.679***	37.927***
	(2.160)	(2.143)	(2.125)
gender	-24.150***	19.225	-8.586***
	(2.262)	(13.955)	(1.956)
repeat_student	-55.793***	-47.263***	-44.287***
	(3.567)	(4.991)	(3.856)
native	7.385*	8.741**	7.144*
	(4.466)	(4.264)	(3.971)
higher_educ_father	14.217**	10.825	11.226*
	(6.658)	(6.682)	(6.498)
lower_educ_father	-20.757***	-15.221***	-22.440***
	(5.464)	(5.330)	(5.915)
native_mother	-0.612	-1.879	1.523
	(3.526)	(3.143)	(2.973)
native_father	6.863*	5.294	4.945
	(3.707)	(4.434)	(3.300)
classtime_math_minutes	-0.029	-0.017	-0.002
	(0.026)	(0.029)	(0.024)
classtime_read_minutes	0.038	0.024	0.001
	(0.029)	(0.029)	(0.027)
classtime_science_minutes	0.045***	0.047***	0.034***
	(0.012)	(0.013)	(0.011)
Year (2009=1)	-17.609**	-14.686**	-12.002*
	(7.187)	(7.114)	(6.668)
2009 * higher_educ_father	11.108	10.930	11.998
	(8.206)	(8.568)	(8.234)
2009 * lower_educ_father	2.419	0.437	4.627
	(6.777)	(6.808)	(6.809)
Year (2018=1)	-23.491***	-19.696***	-24.455***
	(6.157)	(5.781)	(6.829)
2018 * higher_educ_father	-8.657	-4.702	-3.997
	(8.016)	(7.900)	(8.195)

2018 * lower_educ_father	10.136	3.513	7.969
	(6.962)	(6.477)	(7.411)
Constant	137.608***	150.422***	168.980***
	(20.887)	(22.869)	(21.670)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1