CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Does the Type of School Financing Impact Student Performance? - International Comparison based on PISA Dataset

Bachelor's thesis

Author: Emily Slorup

Study program: Economics and Finance

Supervisor: Matej Opatrny MSc., Ph.D.

Year of defense: 2024

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, July 30, 2024

Emily Slorup

Abstract

This thesis examines the influence of government funding on secondary school performance in 65 countries and regions using a fixed-effects approach in the regression models. The three school-level datasets each correspond to one of the dependent variables: mathematics, science and reading. The school-level datasets each include all the relevant variables, meaning the respective dependent variable and explanatory variables. This analysis is based on PISA data that is collected every three years, covering the period from 2006 to 2018. The results revealed a diminishing effect between the percentage of government funding and mathematics, science and reading test scores. The relationship is also statistically significant between the percentage of government funding and the dependent variables. The study also addresses the limited generalisability of this thesis by carefully interpreting the results. Lastly, this thesis prompts researchers to investigate the effects of mixed (private and public) funding of schools on secondary school performance.

JEL Classification	C21, C23, H75, I22, I28
Keywords	secondary school performance, educational determinants, regression analysis, school resources, government funding
Title	Does the Type of School Financing Impact Student Performance? - International Comparison based on PISA Dataset

Abstrakt

Tato práce zkoumá vliv vládního financování na výsledky středních škol v 65 zemích a regionech s využitím přístupu fixních efektů v regresních modelech. Tři soubory dat na úrovni škol odpovídají vždy jedné ze závislých proměnných: matematice, přírodním vědám a čtení. Každý z datových souborů na úrovni školy zahrnuje všechny relevantní proměnné, tedy příslušnou závislou proměnnou a vysvětlující proměnné. Tato analýza vychází z údajů PISA, které se sbírají každé tři roky a pokrývají období od roku 2006 do roku 2018. Výsledky odhalily klesající vliv mezi procentem financování ze strany státu a výsledky testů v matematice, přírodních vědách a čtení. Vztah je rovněž statisticky významný mezi procentem vládního financování a závislými proměnnými. Studie se rovněž zabývá omezenou zobecnitelností této práce, a to prostřednictvím pečlivé interpretace výsledků. Závěrem tato práce podněcuje výzkumníky ke zkoumání vlivu smíšeného (soukromého a veřejného) financování škol na výsledky středních škol.

Klasifikace JEL C21, C23, H75, I22, I28
Klíčová slova výsledky středních škol, determinanty vzdělávání, regresní analýza, školní zdroje, státní financování

Název práce Má druh financování školy vliv na výsledky žáků? - Mezinárodní srovnání na základě souboru dat PISA

Acknowledgments

My appreciation goes out to Matej Opatrny MSc., Ph.D.. Their advice, criticism, and encouragement were really helpful in forming the final version of my bachelor's thesis. My sincere thanks go out to all of my university colleagues and friends, without whom I would not be at this stage. I would like to express my heartfelt thanks to my great aunt and uncle for supporting me, as well as my parents and my whole family for their encouragement. A special thanks to Anna Nakipova for providing endless support.

Typeset in LaTeXusing the IES Thesis Template. Text fine-tuned and grammar-checked by Grammarly.

Bibliographic Record

Slorup, Emily: *Does the Type of School Financing Impact Student Performance? - International Comparison based on PISA Dataset*. Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 52. Advisor: Matej Opatrny MSc., Ph.D.

Contents

List of Tables **viii**

Acronyms **ix**

1 Introduction **1**

2 Literature Review **3**

2.1 Effects of Private and Public Schooling on Secondary School Performance . . . 3
2.2 Effects of School Resources on Secondary School Performance . 4
2.2.1 Other Factors Influencing Secondary School Performance 6
2.3 Methodology Utilised In Relevant Studies . . . 7
2.4 Effects of PISA research . . . 7
2.5 Hypothesis . . . 8

3 Data **9**

3.1 Data . . . 9
3.2 PISA . . . 9
3.3 Dependent Variables . . . 10
3.4 Independent Variables . . . 11
3.4.1 Student Characteristics . . . 11
3.4.2 School Characteristics . . . 13
3.5 Outliers . . . 16

4 Methodology **17**

4.1 Ordinary Least Squares Model . . . 17
4.2 Fixed Effects Model . . . 18
4.3 Model Specifications . . . 19

5 Results **20**
5.1 Fixed Effects Results 20
5.1.1 Mathematics as the Dependent Variable 22
5.1.2 Science as the Dependent Variable 25
5.1.3 Reading as the Dependent Variable 28
5.2 Discussion 31

6 Conclusion **34**

Bibliography **40**

A Results **I**

List of Tables

3.1 Summary Statistics For The Dependent Variables 11
3.2 ISCED Levels and Corresponding Education Levels 13
3.3 Continuous and Dummy Student Characteristics: Descriptive Statistics for Mathematics 13
3.4 Continuous and Dummy Student Characteristics: Descriptive Statistics for Science 14
3.5 Continuous and Dummy Student Characteristics: Descriptive Statistics for Reading 14
3.6 Continuous School characteristics: Descriptive Statistics of Variables for Mathematics 15
3.7 Continuous School characteristics: Descriptive Statistics of Variables for Science 16
3.8 Continuous School characteristics: Descriptive Statistics of Variables for Reading 16

5.1 Fixed Effects: Model Diagnostics 21
5.2 Regression Results: Mathematics 23
5.3 Regression Results: Science 26
5.4 Regression Results: Reading 29

A.1 Random Effects Regression Results I
A.2 Hausman Test Results II
A.3 Variance Inflation Factor (VIF) Values for Each Variable II
A.4 AIC (Akaike Information Criterion) Values for Each Variable . . II
A.5 Full Variable Names for Abbreviations III

Acronyms

FE Fixed Effects

RE Random Effects

OLS Ordinary Least Squares

PISA Program for International Student Assessment

TIMSS Trends In International Mathematics and Science Study

Chapter 1

Introduction

Researchers such as Kapur (2018) and Voyer & Voyer (2014) have been exploring the determinants of education. This debate has been ongoing for decades. This thesis aims to analyse the effect of government funding on secondary school performance based on the Program for International Student Assessment (PISA) test scores in mathematics, science and reading. Based on the author's knowledge, this study is unique in that it explores government funding as the variable of interest in international panel data over a period of time, which has not been explored with PISA data before. This may also be because many studies exploring PISA data analyse a specific country or region, as demonstrated by Giambona & Porcu (2018) and Esen & Adıgüzel (2023).

The literature on this topic provides mixed answers to the impact of government funding on secondary school performance as stated by Martin (2020). Given the mixed findings in the literature, this study presents the following hypothesis: An increase in government spending will result in an increase in PISA test scores. The hypothesis is supported by the papers by Lubienski & Lubienski (2013), Mackenzie (2006), and Sisungo *et al.* (2014). The papers support the notion that government funding positively impacts secondary school performance, and these findings will be explored further in the literature review.

The variety of educational systems makes it hard to establish one universal approach to improving or evaluating secondary school performance. Datasets such as TIMSS (Trends In International Mathematics and Science Study) and PISA (Program for International Assessment) created new opportunities for researchers to understand the various educational determinants that influence secondary school performance across the globe. These sources of data prompted waves of research on secondary school performance (Sjøberg & Jenkins (2022)). While parents and schools could claim that government funding helps improve secondary school performance to receive more funding for more educational resources, others could claim that the funding is ineffective and should be directed towards entirely different areas. It is to be noted that in this study, the following words will be used interchangeably: test scores, school performance and secondary school performance.

Our three datasets each comprise approximately 50,000 observations of secondary schools for each of the dependent variables - mathematics, science, and reading. We use PISA test scores that were conducted by the Organization for Economic Co-operation Development (OECD) as our source of data. In addition to including other school characteristics and student characteristics, the government funding variable is included in the PISA data over the period of 2006 to 2018 as an independent variable. The analysis was carried out using the fixed effects regression models.

The study revealed a statistically significant relationship between government funding and test scores in mathematics, science, and reading. The analysis also revealed a diminishing and non-linear effect of government funding on the mathematics, science and reading scores. Specifically, as the percentage of government funding increases, the student's secondary school performance in these subjects is positively impacted up to a certain point. However, the effect diminishes as the percentage of government funding grows, suggesting that while initial increases in funding have a substantial positive impact on test scores, the benefits of additional percentages of government funding become less pronounced beyond a certain point.

The thesis uses the following structure: Chapter 2 provides the literature review, which examines the related research. Chapter 3 presents the PISA data used in our models and provides details on the PISA data and variables included. Chapter 4 delves into the methodology, outlining the analytical approach and the methods employed. Chapter 5 shows and discusses the results of the regression models. Chapter 6 contains the conclusion, which summarises the key findings of the study and acknowledges its limitations.

Chapter 2

Literature Review

This chapter seeks to examine the existing findings on the influence of government funding on secondary school performance in 65 countries and regions. Additionally, this chapter comprehensively reviews existing research and critically analyses these studies. This chapter will explore the potential policy implications of the current findings and suggest how the findings of this study can be used to influence these policies. It is worth noting that previous research in this area has employed various terms, such as educational achievement and academic performance. The terms are used interchangeably in the field; however, to ensure consistency and clarity in this study, the term secondary school performance will be used throughout.

2.1 Effects of Private and Public Schooling on Secondary School Performance

According to Gross (2017), ever since the 1870s, the question of whether to attend a private or public school has been under debate in America. With papers such as Bedi & Garg (2000) arguing that there are an increasing number of papers claiming that government-provided education is inefficient, governments can look towards private education instead. A study by Suna *et al.* (2020) illustrated how students attending private schools that were socioeconomically stronger scored higher in tests that assessed language, mathematics, and science. The study also acknowledged that when controlling for socioeconomic background, the mean scores of private school students fell dramatically, illustrating the importance of socioeconomic factors when it comes to school performance. Another study by Andrabi *et al.* (2010) corroborates the evidence by showing that students attending the private high school in question obtained significantly higher scores on tests compared to students attending the public high school in Pakistan.

Yet, the assumption that private school students will invariably outperform their public school counterparts is overly simplistic. A paper written by Coleman *et al.* (1982) claimed that regarding secondary schools, private schools outshine public schools specifically in terms of cognitive outcomes. Shortly thereafter, a study by Alexander & Pallas (1983) went on to critique the paper by Coleman *et al.* (1982), claiming that the selection bias was not dealt with properly and the paper should not be used for private school policies. This illustrates how policymakers should be careful when making evidence-based policies. Considering that the government mainly funds education in the majority of developing countries (Bedi & Garg (2000)), it can put pressure on governments to help improve the school performance of students in the best way possible and might disapprove of studies such as Coleman *et al.* (1982).

On the other hand, papers by Lubienski & Lubienski (2013) and Goldhaber (1996) showed that private high schools do not outperform public high schools and the papers strongly advocated for public schools. Another study by Grimes (1994) stated that according to its results, public schools have better teaching of the subject of economics than private schools do. The paper suggests that public schools can provide quality teaching and justifies the ongoing state funding of schools.

2.2 Effects of School Resources on Secondary School Performance

For a long time, researchers, namely Morgan & Sirageldin (1968), have been attempting to investigate whether investing in school resources increases the return on investment or whether the resources are inefficiently allocated, meaning that money is unnecessarily spent. The study by Morgan & Sirageldin (1968) was conducted already in 1968 and examined the positive association between state spending and people's later earnings.

Naturally, when large budgets of money are being used towards public schools, for example, in America, the public becomes concerned about whether it's a valuable investment (Couch *et al.* (1993)). A study by Coleman (1968) from 1966 was amongst the first to research the relationship between government funding and school performance and claimed that it was not a significant relationship. The paper was a groundbreaking study that was funded by the US Congress, and it proves how governments are very eager to see how their funding affects the performance of students. Hanushek (1997) states that the relationship is complex and that it is unrealistic to expect school performance to improve by only redistributing or increasing funding.

A consensus on the relationship between the funding of schools and students' school performance is lacking (Kushebayev & Nygymetov (2022); Martin (2020)). Kushebayev & Nygymetov (2022) also acknowledges that it could not be the government funding that is the issue but rather the ineffective use of the government funding.

On the one hand, studies, namely Mackenzie (2006) and Sisungo *et al.* (2014), reveal a significant positive relationship between the percentage of government funding and secondary school performance. Sisungo *et al.* (2014) also claimed that students in schools in Kenya with 30 per cent of government funding and below performed terribly. However, it should be noted that schools in Kenya are not adequately funded. The paper illustrated that a key level of government funding is needed for students to perform well in secondary school. On the other hand, Garen & Bray (1990) found merely a minor association between government funding and test scores. An example of another study is Kushebayev & Nygymetov (2022), which concludes that a strong statistical link was lacking between secondary school performance and state (government) funding.

Shero & Hart (2022) outlines how new research has moved towards more direct quasi-experimental research, leading to more concise results of a positive relationship between funding and secondary school performance. Cross-sectional studies and correlation analysis are easier to carry out than quantile analysis. The quantile model does not inherently control for individual-specific effects as the fixed effects model. It is to be noted that the quantile regression approach is similar to the fixed effects approach, however they serve different purposes. Quantile regression estimates the conditional quantiles of the dependent variable, whereas the fixed model controls for unobserved heterogeneity in panel data. The fixed effects model is also more sensitive to outliers. Bedi & Garg (2000) asserts that empirical evidence validates test scores as a short-term indicator of educational efficacy.

However, researchers could be looking in the wrong direction. Jimenez & Cox (1989) question the extent to which student characteristics impact school performance over than the school characteristics. This will be further elaborated on in the following section.

2.2.1 Other Factors Influencing Secondary School Performance

Researchers are still discussing the determinants and factors influencing secondary school performance. These papers have been examining educational determinants already from primary school (Li & Qiu (2018)) up to university level (Azhar *et al.* (2014)). When investigating the influence of school funding on secondary school performance, researchers must include additional explanatory variables as well as variables of interest. This increases the model's explanatory power. As the model in this study contains explanatory variables, this subsection will explore frequently employed explanatory variables in education research.

Many studies aim to analyse a large range of relevant variables within their datasets, including the study by Kapur (2018). The study by Kapur (2018) discusses the significance of factors such as learning materials and teachers on secondary school performance in India. The author emphasises the importance of teachers and shows how secondary school performance is dependent on the teacher-to-student relationship. Furthermore, Idris *et al.* (2020) supports the notion that parental education improves secondary school performance. Moreover, a study by Ibrahim *et al.* (2017) demonstrated that parental encouragement and school environment significantly positively affected the secondary school performance of senior secondary girls.

Gender (student characteristic) is a commonly used explanatory variable. In the context of secondary school performance, there is a consensus that females have an advantage in terms of school marks, as shown in the following study. The study by Voyer & Voyer (2014) presented a meta-analysis, which highlighted that female students received higher secondary school performance scores. A recent study by Bentley & Sieben (2022) concluded that in Australia, females have been outperforming males in English and Mathematics. Cornell *et al.* (2016) used the variables, disciplinary structure and student support, to analyse their impact on secondary school performance. Esen & Adıgüzel (2023) used the variable ESCS to investigate the relationship between ESCS and secondary school performance in Turkey using 2018 PISA data. Additionally, the study found that the variable of interest was able to explain greater variance in secondary school performance, specifically for the private schools school.

Giambona & Porcu (2018) claims that research illustrates a positive effect of school size on primary schools but finds the evidence for secondary schools to be inconclusive. Additionally, the study by Giambona & Porcu (2018) found an inverse U-shaped relationship existed between school size and secondary school performance in Italy in 2012 using PISA data. The paper illustrated that secondary school performance improves as school size increases up to a certain threshold. A recent study by Opatrny *et al.* (2023) carried out a meta-analysis on another school-related characteristic: class size. The study concluded that the effect of class size on secondary school performance is zero. Lastly, peer effects appear to have a strong non-linear effect on secondary school performance (Ding & Lehrer (2007); Burke & Sass (2013)).

2.3 Methodology Utilised In Relevant Studies

Given the panel structure of our data, a fixed effects model is appropriate (Wooldridge (2010)). A study by Hanushek *et al.* (2013) employed the fixed-effects approach to school autonomy and its effect on secondary school performance. Hanushek *et al.* (2013) found that the effect of autonomy is positive on secondary school performance in developed and high-achieving countries and vice versa. Another example can be the study written by Teltemann & Schunck (2016), which examined the impact of school segregation on secondary school performance through three waves of PISA country-level data.

2.4 Effects of PISA research

As stated by Schleicher (2017), the role of education is becoming more and more crucial for success in the world economy. Naturally, as noted by Wibowo & Rukayah (2020), countries strive to provide high-quality education already from primary school and onwards. Therefore, many studies have taken the initiative to investigate education based on PISA and TIMSS data, as mentioned in

Sjøberg & Jenkins (2022). PISA began its research in 2000 and continues to this day to facilitate educational improvement. Schleicher (2017) sheds light on the important impact that PISA data brings with OECD's valuable initiative to conduct international assessments and the suggestions it poses to improve educational outcomes. It allows policymakers to develop effective strategies by enabling them to look inward at their own school systems. However, it is worth noting that there is a risk in policymakers using PISA data selectively to appear well in the spotlight of the public debate. Additionally, there is a risk of PISA data fostering competition rather than cooperation between countries.

Williams (2021) addresses the issue with the majority of the research on funding in education. The research considers only individual schools, regions or countries and, therefore, does not take international data into account, which limits generalisability. This perspective aligns with the prevailing trend in PISA-based research, where the focus is on educational determinants within specific regions and individual assessment years. For example, a study by Haw *et al.* (2021) outlines the positive effect of supportive teaching on reading scores recorded in the Phillipness in 2018. According to the author's understanding, valuable research has explored determinants of secondary school performance using PISA data in various contexts; however, a gap persists in our understanding of the influence of school type and public (government) or private funding on secondary school performance on a global scale over a longer period of time.

2.5 Hypothesis

The introduction and literature review sections have explored the necessary studies and literature, and thus, the hypothesis could be formed.

> Hypothesis: An increase in government spending will lead to an increase in PISA test scores

This hypothesis was supported by papers such as Mackenzie (2006), which outline how government funding has a significant impact on secondary school performance. The research aims to contribute to the field by trying to establish a more definitive understanding of how government funding affects secondary school performance.

Chapter 3

Data

3.1 Data

In this chapter, the data set is presented, along with its characteristics. Cross-sectional school-level data was formed from the following source: the Program for International Student Assessment (PISA), which the OECD is responsible for. The three school-level data sets contain 53534, 53534 and 56924 observations of secondary schools for the 5 years (2006, 2009, 2012, 2015 and 2018) from 79 countries and regions for mathematics, science and reading, respectively. It is a diverse spectrum of international data that broadens the applicability and generality of the findings.

The datasets were extracted using an R package. The student dataset was merged with the school dataset based on their common column (School ID) to create a combined dataset. Student data was reported by each student with their school ID and unique ID, whilst the school data included school-specific information (school size, school funding, student-teacher ratio...). Next, the data was aggregated at the school level to create the school-level panel datasets. First, the dependent variables will be introduced, followed by the independent variables. The variables are arranged according to their data type and will be accompanied by their descriptive statistics.

3.2 PISA

According to the PISA Technical Report 2018 (OECD (2018)), PISA collects data and performs standardised assessments every three years (except for 2022, when it implemented an assessment four years later), which is why there is no possibility of comparing results across all years. PISA tests are standardised assessments carried out in every included country and region to examine the reading, mathematics and science literacy in 15-year-old pupils to improve educational outcomes and help transform education. Every chosen year, one of the three literacies is picked to be studied in detail, compared with the other two assessments being studied in less detail as other focus areas. Reading literacy was studied in detail in 2000, followed by mathematics in 2003, and science at the end of the cycle in 2006. This cycle continues to repeat.

As stated in Hanushek *et al.* (2013), the PISA sample is stratified in two stages; schools are first randomly picked from each country and region, followed by students being randomly chosen within each school. Furthermore, the standardised tests are performed using paper and pencils. With similar test environments and data collection methods, the international data is comparable between countries and regions. All variables were measured according to the same scale and indexed in a consistent format across all countries and regions. Student weights were applied to each variable to ensure the whole population was captured. Despite the aggregation of the data reducing variation, it was necessary to be able to compare changes in the variables over time.

3.3 Dependent Variables

As mentioned previously, the subject of mathematics, science or reading was tested in detail during specific years according to the cycle; however, data was collected in all three areas on the general level. This is also the reason why the general mathematics, science and reading scores were selected to be the dependent variables to compare the variables over time. The dependent variables each represent a result from PISA tests and thus are numeric variables. The range of the variable is between 0 and 1000. For the purpose of this study, all three areas will be examined to allow for in-depth and focused research, as the three areas are main subjects in many countries as well as regions and are compulsory for all PISA test-taking students. According to Giambona & Porcu (2018), the test results are an appropriate quantitative measure of evaluating secondary school performance. In this study, it is worth noting that each student writes the standardised assessments for all three subjects. The descriptive statistics of the variables are displayed in Table 3.1 below. The word location is used to represent the country or region in the table. The three countries or regions with the highest scores are also displayed in Table 3.1 for each subject. As can be seen in Table 3.1 below, Singapore, Hong Kong, and Finland are the top three countries or regions in all three subjects. In the context of an international dataset, it can be useful to highlight the leading countries and regions.

Table 3.1: Summary Statistics For The Dependent Variables

Subject	**Location**	**Min**	**1st Qu.**	**Median**	**Mean**	**3rd Qu.**	**Max**
Math	All Countries	98.23	405.7	467.6	463.3	518.0	812.5
	Hong Kong	345.6	507.4	558.7	553.5	602.5	690.5
	Finland	314.2	507.5	528.4	528.1	551.8	715.8
	Singapore	379.7	517.2	547.1	555.7	584.5	722.5
Science	All Countries	20.18	415.5	467.2	470.6	525.2	787.6
	Hong Kong	322.4	501.6	545.3	540.6	585.3	661.8
	Finland	311.2	525.3	546.9	545.1	570.1	745.6
	Singapore	349.7	497.8	531.8	538.4	568.3	725.3
Reading	All Countries	0.122	405.1	466.3	460.0	516.9	737.5
	Hong Kong	397.1	492.5	539.6	534.9	579.7	662.0
	Finland	271.9	511.4	532.0	530.7	554.8	715.2
	Singapore	337.7	483.4	514.2	523.6	553.6	700.2

Note: The table presents data for all 65 countries and the top three performers in the datasets. The number of observations was 53532, 53532 and 56924 for the math, science and reading variables, respectively. Summary statistics were calculated using R.

3.4 Independent Variables

3.4.1 Student Characteristics

As previously mentioned, PISA collects data on each pupil. Due to the nature of the variables, they were placed into two categories: continuous and dummy variables. The descriptions for each variable are derived from the OECD PISA Technical Report 2022 and the OECD PISA DATA Explorer. The purpose of including the student characteristics was to investigate if and how the variables have an impact on secondary school performance.

Continuous variables

One student characteristic is the ESCS, an index of economic, social, and cultural status. It is worth noting that it is an index, and therefore, positive values demonstrate an above-average status and negative values demonstrate a below-average status. The index consists of categories such as home possessions, the number of books at home and food insecurity. The variable was supported by a study by Esen & Adıgüzel (2023), which argued for the critical impact of ESCS on the secondary school performance of students. It is expected that the three test scores will be higher with a higher ESCS score. This relevant variable was selected to examine whether a student's background would influence the dependent variable. Tables 3.3, 3.4 and 3.5 describe the summary statistics of the regression variable for each dependent variable.

Dummy variables

The following variables further describe students in detail, which is vital in determining secondary school performance from the perspective of a student's background. After careful consideration, the Male variable, which represents the proportion of males in a school, was selected. As previously mentioned in Chapter 2, there has been a consensus amongst researchers that females perform better in academics than males. The research suggests that the male variable should influence the dependent variables negatively. Descriptive results are stated in Tables 3.3, 3.4, and 3.5.

The variables Computer and Internet represented the proportion of students that had access to a computer (laptop, desktop, or tablet) or internet (Wi-fi and not including smartphones) that the student could use for school work. The variables, Computer and Internet, were selected based on a study by Grace-Martin & Gay (2001), which researched the relationship between browsing activity and a student's final grade. It is expected that with access to the internet and a computer, a student will have more access to better educational resources and be able to perform better in secondary school.

Studies have illustrated that a student's secondary school performance is closely associated with the academics of their parents, as illustrated by Idris *et al.* (2020). The study displayed how a better educational status of the parents had an important and positive relationship with their children's secondary school performance. A similar outcome is expected in this study. The Mother's education and Father's education variables are indexes that represent the mother's and father's level of education (MISCED and FISCED). ISCED stands for the International Standard Classification of Education and is implemented to compare various educational systems.

The mother and father education variables were converted to factors ranging from 1 to 4 as the responses ranged from "less than ISCED1", "ISCED 1", "ISCED 2", "ISCED 3A", and "ISCED 3B, C". As can be seen below, "ISCED 3A" and "ISCED 3B, C" both represent a high school education and were converted to a four. The levels of education are described in Table 3.2 in detail below. The dataset did not contain information on parental education beyond the high school level, which limits the ability to explore such variations in parental education. The data only reflects only their highest attainment within the secondary education and below categories.

ISCED Level	**Education Level**
Less than ISCED 1	Less than primary education (pre-primary or no formal education)
ISCED 1	Primary education (ages 6 to 11)
ISCED 2	Lower secondary education (middle school or junior high school, ages 12 to 15)
ISCED 3A	Upper secondary education with access to higher education (high school, ages 15 to 18m preparation for university)
ISCED 3B, C	Upper secondary education with access to higher education (ages 15 to 18, preparing for vocational training or workforce)

Table 3.2: ISCED Levels and Corresponding Education Levels

Table 3.3: Continuous and Dummy Student Characteristics: Descriptive Statistics for Mathematics

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
ESCS	-4.8514	-0.8226	-0.2323	-0.3399	0.2528	2.3800
Male	0.0000	0.4050	0.4988	0.5062	0.5968	1.0000
Computer	0.0000	0.7200	0.9216	0.8037	1.0000	1.0000
Internet	0.0000	0.7028	0.9422	0.7968	1.0000	1.0000
Mother Educ	1.000	2.675	2.971	2.901	3.163	4.000
Father Educ	1.000	2.667	2.981	2.919	3.201	4.000

Note: 53532 is the number of observations, and R was employed for calculating summary statistics

3.4.2 School Characteristics

The school characteristics picked were all continuous and numerical.

Table 3.4: Continuous and Dummy Student Characteristics: Descriptive Statistics for Science

Variable	**Min**	**1st Qu.**	**Median**	**Mean**	**3rd Qu.**	**Max**
ESCS	-4.8514	-0.7362	-0.1847	-0.2909	0.2773	2.3800
Male	0.0000	0.4045	0.5000	0.5079	0.6000	1.0000
Computer	0.0000	0.7600	0.9323	0.8206	1.0000	1.0000
Internet	0.0000	0.7500	0.9513	0.8126	1.0000	1.0000
Mother Educ	1.000	2.675	2.971	2.901	3.163	4.000
Father Educ	1.000	2.667	2.981	2.919	3.201	4.000

Note: The number of observations is 53532, and the descriptive statistics were computed using R.

Table 3.5: Continuous and Dummy Student Characteristics: Descriptive Statistics for Reading

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
ESCS	-4.8514	-0.8175	-0.2286	-0.3400	0.2521	2.3800
Male	0.0000	0.4047	0.4990	0.5069	0.5985	1.0000
Computer	0.0000	0.7204	0.9235	0.8039	1.0000	1.0000
Internet	0.0000	0.7011	0.9428	0.7960	1.0000	1.0000
Mother Educ	1.000	2.642	2.960	2.882	3.152	4.000
Father Educ	1.000	2.636	2.963	2.902	3.190	4.000

Note: The number of observations is 56924, and the summary statistics were generated through R analysis

Continuous Variables

Four variables make up all of the total funding for a school year and, thus, add up to 100 per cent. The four variables are government funding, student funding, philanthropy funding and other funding. Government funding is the main variable of interest, as this study aims to investigate the importance and influence of government funding on secondary school performance. The government funding depicts the percentage of funding coming from the government (including departments, local, regional, state and national).

According to the PISA OECD Data Explorer, student funding represents the funding coming from student fees or school charges paid by the parents. The PISA OECD Data Explorer defines philanthropy funding as the percentage of funding coming from benefactors, donations, bequests, sponsorships, and parent fundraising. As the philanthropy and student funding variables could be put in the same category, the variables were added together to create a private funding variable. The other funding variable did not specify the exact sources of funding and, therefore, would produce ambiguous results. Through comprehensive research, there are conflicting studies. Many studies, such as Mackenzie (2006), state that government funding helps secondary school performance. Others, such as Garen & Bray (1990), reveal a minor impact of government funding on the secondary school performance of students and suggest that government funding should not be increased over private funding. Hence, the direction of government funding could be favourable or unfavourable.

According to OECD (2018), the student-teacher ratio (stratio) variable depicts the number of enrolled male and female students divided by the total number of full-time and part-time teachers. The impact of a teacher has been seen as invaluable in papers. For example, the variable Stratio in a paper written by Giambona & Porcu (2018) highlighted the positive relationship between stratio and secondary school performance. It is expected that a bigger student-teacher ratio will have a negative effect on the student's secondary school performance as the student receives less of the teacher's attention.

The variable, school size, is the index of total enrolment at school. It is presumed that with a larger school size, the students would have access to more facilities and teachers. However, papers have reported conflicting results. The paper by Giambona & Porcu (2018) states through their study that school size has an inverse U-shaped relationship with secondary school performance. In Tables 3.6, 3.7 and 3.8, the descriptive statistics for the variables are provided. The difference in how the funding variable and science/reading/math scores vary in the descriptive results can be attributed to missing data for some student test scores. In addition, the school-level dataset in relation to the reading variable has a slightly higher number of observations.

Table 3.6: Continuous School characteristics: Descriptive Statistics of Variables for Mathematics

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Fund Public	0.0	75.0	96.0	79.9	100.0	100.0
Fund Private	0.0	0.0	0.0	14.5	12.0	100.0
Stratio	0.018	9.484	12.592	12.886	15.926	27.578
School Size	0.0	296.0	567.0	644.5	913.0	2027.0

Note: Summary statistics for the 53,532 observations were computed using R statistical software.

Table 3.7: Continuous School characteristics: Descriptive Statistics of Variables for Science

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Fund Public	0.0	79.0	97.0	81.48	100.0	100.0
Fund Private	0.0	0.0	1.0	15.65	17.0	100.0
Stratio	0.018	9.367	12.450	12.695	15.681	27.576
School Size	0.0	295.0	564.0	638.6	906.0	1950.0

Note: For the 53,532 observations, the summary statistics were computed using R statistical software.

Table 3.8: Continuous School characteristics: Descriptive Statistics of Variables for Reading

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Fund Public	0.0	76.0	96.0	80.61	100.0	100.0
Fund Private	0.0	0.0	1.0	16.42	20.0	100.0
Stratio	0.018	9.583	12.800	14.370	16.477	723.0
School Size	0.0	304.0	581.0	658.2	935.0	1988.0

Note: The dataset comprised 56,924 observations, and descriptive statistics were generated through R analysis.

3.5 Outliers

Outliers were also removed, as the range of some variables in the summary statistics appeared to be quite large. Histograms were created, and they suggested the presence of outliers. The identification and removal of outliers was done using the interquartile range (IQR) method. This method effectively identifies extreme values that could skew the results. 6032, 6550 and 2730 outliers were removed for the mathematics, science and reading school-level datasets, respectively. This means that 89, 88 and 95 per cent of observations were kept for the mathematics, science and reading school-level datasets, respectively. The data was cleaned to improve the data distribution.

Chapter 4

Methodology

This section outlines the methodology for analysing the dataset. As previously mentioned, the final data sets are three international panel datasets of secondary schools. Firstly, the analysis considers the Ordinary Least Squares (OLS) Model, Random Effects (RE) Model and Fixed Effects (FE) Model. The diagnostics of the model revealed that the FE Model is the better-fitting model. Therefore, the FE model will be used. The model is ideal for when working with aggregated panel data as it controls for unobserved heterogeneity (Wooldridge (2010)). A paper written by Hanushek *et al.* (2013) also carried out a fixed-effects approach on PISA data. The AICc criterion will be used to specify the FE models. OLS analysis provides a valuable perspective as it does not lose individual-level information, complementing the insights gained from the aggregated FE models. Aggregating the data means losing the variations and specific characteristics of each student's secondary school performance across the three subjects. The analysis aims to contribute crucial insights into the relationship between government funding and secondary school performance.

4.1 Ordinary Least Squares Model

The Ordinary Least Squares Model is a standard method used to analyse econometric data. The following OLS regression model was created to estimate the link between the dependent and independent variables:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 \text{fund\_gov}_i + \epsilon_i \tag{4.1}$$

where Y_i is the dependent variable, (β_0) is the intercept, (β_1) is the slope,

X_i is the vector of independent variables, β_2 is the coefficient for the dummy variable fund_gov$_i$, and ϵ_i is the error term.

4.2 Fixed Effects Model

The Fixed Effects Model is a method commonly used with panel data, as it takes unobserved heterogeneity into account. In the context of this study, this method should be appropriate when addressing unobservable variables that remain fixed over time but differ among schools. This particular model should take into account the school's unobserved differences as well as other factors.

Since the FE Model considers time and unit-fixed effects, bias is thereby lowered, and more accurate results can be found. For example, if the variable of interest was the student-teacher ratio's effect on test scores, the estimates could be inaccurate because of unobserved differences among schools if the time and unit-fixed effects are not considered. The following regression model is written below, and it consists of s schools, t time periods and c countries or regions:

$$y_{stc} = \beta_0 + \beta_k X_{stc} + a_{sc} + \lambda_{tc} + u_{stc} \tag{4.2}$$

Where y_{stc} represents the dependent variable for school s at time t in country or region c, X_{stc} represents the vector of k independent variables, a_{sc} represents the unobserved school-specific effect, λ_{tc} represents the time-fixed effect, and u_{stc} represents the idiosyncratic error term.

In the first transformation of the equation, the unobserved school-specific effects (a_{sc}) and time-fixed effects (λ_{tc}) will be removed by subtracting the school's mean over time from each variable:

$$y_{stc} - \bar{y}_{stc} = \underbrace{\beta_0 - \bar{\beta}_0}_{0} + \beta_k(X_{stc} - \bar{X}_{stc}) + \underbrace{a_{sc} - \bar{a}_{sc}}_{0} + \underbrace{\lambda_{tc} - \bar{\lambda}_{tc}}_{0} + u_{stc} - \bar{u}_{stc}. \tag{4.3}$$

where $\bar{y}_{stc}$, $\bar{X}_{stc}$, and $\bar{u}_{stc}$ are the means for school s over t time periods for country or region c. The previous equation will be simplified to reveal the final equation:

$$\ddot{y}_{stc} = \beta_k \ddot{X}_{stc} + \ddot{u}_{stc} \tag{4.4}$$

Where $\ddot{y}_{stc}$, $\ddot{X}_{stc}$, and $\ddot{u}_{stc}$ are the final variables after the within-school (time-demeaned) transformation used to alter the original values illustrated in Equation 4.3.

4.3 Model Specifications

Models with different percentages of government funding have been specified for this study as they might reveal whether there's a threshold effect, meaning that the impact of more government funding might have a stronger or weaker effect at different funding percentages. It also examines how the impact of funding varies depending on the government funding percentage. The first model is the OLS model, which contains all the students from each dataset prior to aggregation. The second model contains all the schools in the school-level datasets, which shows how variability in the range of government funding affects the dependent variables. This model serves as a baseline by including all schools in each school-level dataset. The third model includes at least one per cent up to 25 per cent of government funding to be able to examine only schools that receive less than a quarter of the total funding from the government. The fourth, fifth and sixth models consist of schools that receive more than 25 and up to 50 per cent, more than 50 and up to 75 per cent, more than 75 and up to 100 per cent of government funding. The last model regards exclusively schools, where all of the school funding comes from the government. In other words, it strictly examines the relationship between government funding and secondary school performance. This can provide a nuanced understanding of how effective government funding can be in influencing secondary school performance.

Chapter 5

Results

Within this chapter, the findings of the analysis will be provided.

In this section, the results from the specifications of the fixed-effects models will be presented. As mentioned previously, the model addresses unobserved heterogeneity, which increases the reliability of the results. The model appeared to work well with the data. The details of the analysis are discussed, together with a thorough discussion of their implications. The results of the model diagnostics of the RE Models are shown in Table A.1 in Appendix A.

5.1 Fixed Effects Results

The data shall be analysed using the panel data model below to investigate the relationship between student performance and public funding. With the description and research on the variables included, it is quite certain how the majority of the independent variables will affect the dependent variable. The coefficient β_7 will be the main focus of the regression. Each variable is denoted for school s and country c at time t.

$$
\begin{aligned}
score_{sct} = \beta_1 fundgov_{sct} + \beta_2 stratio_{sct} + \beta_3 school\_size_{sct} + \\
+ \beta_4 mothereduc_{sct} + \beta_5 fathereduc_{sct} + \beta_6 male_{sct} + \\
\beta_7 computer_{sct} + \beta_8 internet_{sct} + \beta_9 escs_{sct} + \epsilon_{sct}
\end{aligned}
$$

Firstly, it is necessary to select the model that is the best fit for the panel data that is to be estimated. Thus, the Hausman Test (a test frequently used with panel data) was carried out, and the results determined that the Fixed

Effects Model was the superior model to the Random Effects Model. The results are provided in Table A.2 in Appendix A.

Before the regressions are carried out, the assumptions of the FE model need to be tested. The dataset consists of PISA data, which was randomly sampled with students and schools. Thus, the assumption of a random sample holds. The multicollinearity between variables was tested using the vif package in R. There appeared to be no significant correlation between the variables, as is evident in Table A.3 in Appendix A. Therefore, no variables are to be excluded. Considering that the standard errors were clustered and weights were incorporated, the model should be robust and accurate in its findings.

The fourth assumption of the FE model is that strict exogeneity holds. It is difficult to ensure that the error term is entirely unrelated to all explanatory variables. However, weak homogeneity will still hold. The estimator can be asymptotically unbiased due to a sufficiently big sample size. The three school datasets contain over 50,000 observations each; thus, the assumption holds as similarly claimed by Sokolakova (2023).

It is necessary to test the assumptions of heteroskedasticity and serial correlation. The Breusch-Godrey and Studentized Breusch-Pagan tests are commonly used to test for serial correlation and heteroskedasticity, respectively. The results in Table 5.1 indicate they are both present in the models. Thus, HAC robust standard errors and the clustering of robust standard errors and weights are implemented to address the presence of heteroskedasticity and serial correlation and provide more reliable estimates. The coefficients will have to be interpreted carefully. A similar method was carried out by ČEJKA (2023), who wrote a paper using a fixed-effects approach and implemented these methods.

Table 5.1: Fixed Effects: Model Diagnostics

Subject	S. BP Test Stat.	BG Test Stat.	p-value
Math	938.15	3573.2	$< 2.2 \times 10^{-16}$
Science	970.85	2372.2	$< 2.2 \times 10^{-16}$
Read	934.54	1907.7	$< 2.2 \times 10^{-16}$

Due to the violations of the fifth and sixth assumptions of the fixed effects model, it can be concluded that the fixed effects estimator is not BEST LINEAR UNBIASED ESTIMATOR (BLUE). Additionally, heteroskedasticity and serial correlation have to be addressed using robust estimation; the estimator is efficient but not consistent.

It is apparent in Table A.4 in Appendix A that the AICc value of the model with the additional variable of Stratio was higher than those of the model with one variable only. This higher AICc value suggests that the inclusion of Stratio did not enhance the model's fit sufficiently to justify its complexity and did not improve the model's ability to explain the variability in the dependent variable. Therefore, the Stratio variable was excluded to use the highest quality model. A paper by Chiatante *et al.* (2021) illustrated how independent variables were selected according to the AICc criterion. Consequently, the subsequent model is to be carried out:

$$\begin{aligned} score_{sct} = {} & \beta_1 fundgov_{sct} + \beta_2 school\_size_{sct} + \\ & + \beta_3 mothereduc_{sct} + \beta_4 fathereduc_{sct} + \beta_5 male_{sct} + \\ & + \beta_6 computer_{sct} + \beta_7 internet_{sct} + \beta_8 escs_{sct} + \epsilon_{sct} \end{aligned}$$

5.1.1 Mathematics as the Dependent Variable

Table 5.2 displays OLS and fixed effects estimations of the seven models for mathematics as the dependent variable. The F statistics demonstrate that the models are highly statistically significant and reveal that at least one of the independent variables is significantly related to the dependent variable. However, the adjusted R^2 statistics of 0.20 are relatively low; however, that is not surprising since there is a limitation of variables, and therefore, the adjusted R^2 statistics fall in a plausible range. In other words, the model only explains approximately 20 per cent of the variability.

All variables in the OLS model are statistically significant, whereas the FE models demonstrate that the mother education variable was statistically insignificant. The variables school size, mother education, father education, internet and ESCS show an effect that was approximately half as strong as that observed in the FE model. The variable computer demonstrates the opposite effect, as its effect was roughly twice as strong as that illustrated in the FE model. The impact of government funding on the dependent variable is consistent across the OLS model and Model 2. Surprisingly, the Male variable demonstrates a positive effect in contrast with the FE models.

Table 5.2: Regression Results: Mathematics

Variable	**OLS**	**(0-100)**	**(1)**	**(25)**	**(50)**	**(75)**	**(100)**
Intercept	429.6***	-	-	-	-	-	-
	(0.4685)	-	-	-	-	-	-
fund_gov	0.0628***	0.0877***	7.686***	9.881***	7.327***	0.728	-5.624***
	(0.0025)	(0.0107)	(1.662)	(1.408)	(1.020)	(0.620)	(1.094)
school_size	0.0133***	0.0217***	0.0215***	0.0214***	0.0213***	0.0216***	0.0212***
	(0.0001)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
male	8.515***	-11.838***	-11.802***	-11.839***	-11.845***	-11.883***	-11.733***
	(0.1506)	(1.327)	(1.328)	(1.328)	(1.328)	(1.329)	(1.328)
mother_educ	0.855***	1.682	1.103	1.036	1.194	1.263	1.161
	(0.0997)	(1.145)	(1.144)	(1.144)	(1.144)	(1.144)	(1.144)
father_educ	1.541***	5.954***	5.374***	5.342***	5.112***	5.476***	5.415***
	(0.0973)	(1.114)	(1.113)	(1.112)	(1.113)	(1.113)	(1.113)
computer	26.62***	14.488***	16.156***	16.723***	15.633***	15.583***	15.295***
	(0.2576)	(2.560)	(2.558)	(2.559)	(2.556)	(2.563)	(2.558)
internet	22.72***	47.642***	48.703***	48.191***	48.511***	48.517***	49.065***
	(0.2527)	(2.181)	(2.180)	(2.179)	(2.179)	(2.180)	(2.182)
escs	28.16***	48.225***	47.066***	46.957***	47.128***	47.303***	47.092***
	(0.08393)	(0.621)	(0.610)	(0.610)	(0.609)	(0.614)	(0.610)
R^2	0.1946	0.47187	0.47119	0.47161	0.47164	0.47089	0.47126
Adj. R^2	0.1946	0.20168	0.20066	0.20129	0.20134	0.20021	0.20078
F-statistic	40690	3955.12	3944.4	3950.99	3951.54	3939.67	3945.59
Num. obs	1515328	53532	53532	53532	53532	53532	53532
Num. schools	18110	50135	1804	2531	4872	33443	4329

Note: The full variable names are listed in Table A.5
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Referring to our variable of interest (Government Funding), the variable is highly statistically significant across five models (except for model 6) at a 1 per cent significance level, signifying signs of effect. Model 2 (0-100) illustrates that for every additional percentage of government funding a school gains, the math score is expected to rise by 0.0877 points. Model 3 (1) demonstrates that for schools receiving government funding between 1 and 25 per cent, a one per cent increase in government funding would increase the mathematics score by 7.686 at a 1 per cent significance level. It also partially accepts the hypothesis of this study, which will be further elaborated on in this section. Models 3 to 7 illustrate that there is a mostly positive effect; however, it is a diminishing effect of the influence of government funding on mathematics scores, as the effect ranges from 9.881 to -5.624 points. In contrast with the other variables, such as ESCS, the magnitude of the government funding variable is very small. This aligns with research by Garen & Bray (1990) claiming that there is only a small association between government funding and secondary school performance.

Thus, a non-linear relationship between government funding and secondary school performance appears to be present. This could result from the allocation of resources, as it could be more difficult to allocate resources efficiently as funding rises. Additionally, the results suggest that both private and public funding could be beneficial to help schools effectively, and it could imply that a mix of both should be utilised. The research opens a new path to examine mixed funding for schools. Lastly, it sheds light on the idea that government funding does not automatically lead to better educational outcomes and emphasises the complexity of the issue. It also aligns with research that does not support the claim that secondary school performance will improve by simply increasing the proportion of government funding.

At the 0.01 level of statistical significance, the school size variable has a slight positive impact. The mathematics score would increase by roughly 0.02 points for the FE models as school size increased. This could be because attending a larger school would give students a higher availability of resources. The male variable, which is a highly significant variable at a 1 per cent significance level, demonstrates that, on average, a male would score approximately 12 points less than their female counterparts for the FE models. This finding supports the notion that females generally perform better in a school setting.

The results suggest a positive association between a reported mother's education and their children's mathematics scores. The estimated effect is approximately 1. However, this correlation may be inflated as a result of the variable's limited range. This effect is statistically insignificant for all the FE models, indicating that the evidence is weak. The practical significance of such a small increase in the context of the total score is limited, making it challenging to draw meaningful conclusions from this variable.

The father's education variable, on the other hand, is highly statistically significant across all FE models at a significance level of 1 per cent, with a consistently positive effect ranging from 5.112 to 5.954 points. The positive coefficient could be due to parents being responsible for creating a stimulating learning environment at home to foster excellent secondary school performance in their children. However, the estimated effect might not fully capture the impact of a higher father's education.

The computer variable is statistically significant at the 1 per cent significance level, indicating that access to a computer is associated with a significant increase in mathematics scores, ranging from 14.488 to 16.723 points for the FE models. This result underscores the importance of technological resources in enhancing student's learning experiences and improving their secondary school performance.

Similarly, the internet variable shows a significant positive effect at a 1 per cent significance level on mathematics scores, with an increase ranging from 47.642 to 49.065 points. This follows the logical reasoning that access to resources can help improve the grasp of concepts and understanding of the content. The ESCS variable is also strong, positive, and statistically significant across all models. For example, an increase of 1 unit in ESCS corresponds to an approximate 47-point increase in mathematics points. This could be due to having greater access to resources, such as books and study materials, in a perhaps more supported environment.

5.1.2 Science as the Dependent Variable

Table 5.3 displays OLS and fixed effects estimations of the seven models for science as the dependent variable. The F-test results demonstrate strong statistical significance. The models account for 22 per cent of the overall variability in the science variable.

Table 5.3: Regression Results: Science

Variable	**OLS**	**(0-100)**	**(1)**	**(25)**	**(50)**	**(75)**	**(100)**
Intercept	442.1***	-	-	-	-	-	-
	(0.4646)	-	-	-	-	-	-
fund_gov	0.0472***	0.0748***	4.959**	10.267***	6.819***	1.562**	-7.271***
	(0.0024)	(0.0104)	(1.606)	(1.360)	(0.986)	(0.599)	(1.057)
school_size	0.0112***	0.0204***	0.0202***	0.0201***	0.0200***	0.0203***	0.0198***
	(0.0001)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
male	-0.469**	-26.388***	-26.370***	-26.384***	-26.392***	-26.450***	-26.240***
	(0.1493)	(1.283)	(1.284)	(1.283)	(1.283)	(1.284)	(1.283)
mother_educ	0.804***	1.104	1.571	1.693	1.524	1.426	1.582
	(0.0988)	(1.106)	(1.106)	(1.105)	(1.105)	(1.106)	(1.105)
father_educ	1.899***	8.211***	7.733***	7.669***	7.467***	7.837***	7.736***
	(0.0965)	(1.076)	(1.076)	(1.075)	(1.076)	(1.076)	(1.075)
computer	26.51***	27.080***	28.425***	29.160***	28.044***	27.749***	27.551***
	(0.2554)	(2.474)	(2.472)	(2.473)	(2.470)	(2.477)	(2.471)
internet	20.12***	42.528***	43.391***	42.938***	43.269***	43.291***	43.987***
	(0.2506)	(2.108)	(2.107)	(2.106)	(2.106)	(2.107)	(2.108)
escs	28.00***	45.451***	44.494***	44.322***	44.509***	44.781***	44.431***
	(0.08322)	(0.601)	(0.590)	(0.589)	(0.588)	(0.593)	(0.589)
R^2	0.1837	0.48676	0.48615	0.48684	0.48670	0.48611	0.48670
Adj. R^2	0.1837	0.2242	0.22327	0.22431	0.22411	0.22321	0.2241
F-statistic	37900	4198.42	4188.1	4199.64	4197.43	4187.44	4197.3
Num. obs	1515328	53532	53532	53532	53532	53532	53532
Num. schools	18110	50135	1804	2531	4872	33443	4329

Note: The full variable names are listed in Table A.5
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Similarly to the previous OLS model, all variables included are statistically significant at the 1 per cent significance level. The estimated coefficients for government funding and computer remained consistent between the OLS model and Model 2 (FE model). The OLS model demonstrates a comparable impact with the computer and mother education variables. The school size, internet and ESCS variables exhibit a smaller impact on the dependent variable when compared with Model 2 (FE model). The male variable also illustrates a negative effect, and the father's education variable exhibits a positive impact. However, the effect for both variables is minimal.

A diminishing effect in the science variable is apparent across all models, with a significance level of 1 per cent. Between Models 2 and 3, which contain schools with government funding between 1 and 25 per cent, and 26 and 50 per cent, respectively, the effect appears to be positive as the effect on math scores increases from 4.959 to 10.267. As we move across Model 3 to Model 6, which contains schools with government funding between 51 and 75, 76 and 100, and 100 per cent, respectively, the effect appears to reverse, with scores decreasing up to -7.271 points. The results of the mathematics variable produced very similar findings.

The school size variable indicates statistical significance at the 0.01 level, showing a small positive effect of 0.02. At a 1 per cent significance level, the coefficient of the male variable is statistically significant. This means that male students would score approximately 26 points less than female students. The size of the effect is twice as large as the observed effect for the mathematics variable.

The mother's education does not have a statistically significant impact on mathematics scores. When the additional level of education increases (specifically primary or secondary education), there is an approximate increase in the science score by 1.104 to 1.693 points. Similarly to the previous dependent variable, the effect is not present here. At a significance level of 1 per cent, the analysis strongly indicates that a father's education is a statistically significant variable. There is a positive link between each additional ISCED level of education below the university level and an increase in secondary school performance by 7.467 to 8.211 points. These results weakly support the idea that higher parental education boosts better secondary school performance.

At a 0.01 significance level, the use of a Computer has a significant positive impact on the science-dependent variable. Specifically, having access to a computer can be associated with an increase of approximately 28 points in overall science test scores. This is aligned with findings from the previous dependent variable (mathematics), but the effect roughly doubled. At a 1 per cent significance level, access to the internet has a significant impact on the science-dependent variable. In addition, secondary school performance is influenced by having access to the internet. With greater access to the internet, secondary school performance increases by 42.528 to 43.987 points.

Lastly, the ESCS variable has a statistically significant impact of approximately 44 points on science test scores, with a 1 per cent significance level. The result is logical, as with a high ESCS, those students are able to focus on their studies with less stress or uncertainty. There seems to be a close resemblance between the results for mathematics and science with the majority of the variables.

5.1.3 Reading as the Dependent Variable

Table 5.4 displays OLS and fixed effects estimations of the seven models for reading as the dependent variable. According to the F-test results, every model has a high level of statistical significance. Roughly 28 per cent of the total variability in the reading variable can be attributed to the models.

All model predictors are statistically significant in the OLS model. The OLS model illustrates a slightly smaller R^2 in comparison with the FE model. The OLS model yields a similar result to that of the government funding and computer variable of Model 2 (FE model). The impact of the school size, male, mother's education, father's education, internet and ESCS is less pronounced in the OLS model.

At a significance level of one per cent, a declining influence on the reading variable is seen in all models. The effect appears to be favourable between Models 2 and 3, when schools receiving less government funding (50 per cent or less) see an improvement in maths results from 12.507 to 14.574. The impact seems to reverse as we go from Model 3 to Model 6, with higher levels of government funding (50 per cent and higher) and scores dropping as much as -10.574 points.

Table 5.4: Regression Results: Reading

Variable	OLS	(0-100)	(1)	(25)	(50)	(75)	(100)
Intercept	453.5***	-	-	-	-	-	-
	(0.4664)	-	-	-	-	-	-
fund_gov	-0.0201	-0.020*	12.507***	14.574***	8.434***	-1.879**	-10.574***
	(0.0024)	(0.0096)	(1.477)	(1.308)	(0.951)	(0.574)	(1.019)
school_size	0.0096***	0.0209***	0.0207***	0.0207***	0.0207***	0.0209***	0.0203***
	(0.0001)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
male	-32.88***	-65.488***	-65.384***	-65.438***	-65.485***	-65.431***	-65.223***
	(0.1499)	(1.246)	(1.245)	(1.244)	(1.245)	(1.246)	(1.245)
mother_educ	0.449***	1.350	1.245	1.153	1.411	1.385	1.351
	(0.0992)	(1.051)	(1.049)	(1.048)	(1.049)	(1.050)	(1.048)
father_educ	2.257***	10.491***	10.474***	10.439***	10.211***	10.530***	10.497***
	(0.0968)	(1.024)	(1.021)	(1.021)	(1.022)	(1.022)	(1.021)
computer	29.71***	29.907***	30.176***	30.787***	29.475***	30.152***	28.613***
	(0.2564)	(2.331)	(2.327)	(2.327)	(2.326)	(2.333)	(2.328)
internet	21.94***	44.232***	44.429***	43.626***	43.961***	43.977***	45.139***
	(0.2516)	(1.995)	(1.989)	(1.987)	(1.988)	(1.990)	(1.991)
escs	27.41***	44.435***	44.440***	44.299***	44.546***	44.441***	44.409***
	(0.084)	(0.576)	(0.565)	(0.565)	(0.565)	(0.569)	(0.565)
R^2	0.2007	0.51857	0.51942	0.52008	0.51951	0.51865	0.51987
Adj. R^2	0.2007	0.2793	0.28058	0.28157	0.28071	0.27943	0.28126
F-statistic	42270	5119.8	5137.29	5150.94	5139.11	5121.46	5146.63
Num. obs	1515328	56924	56924	56924	56924	56924	56924
Num. schools	18891	52913	2059	2682	5138	35118	4585

Note: The full variable names are listed in Table A.5
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

There was a slight positive impact, with the school size variable displaying statistical significance at the 0.01 significance level. The reading score would increase by roughly 0.02 with an increase in school size. The difference in the effect of school size on reading and the other dependent variables is negligible. Surprisingly, the male variable has the largest effect at a 1 per cent significance level, illustrating a negative impact on the dependent variable. It was revealed in Models 1 to 6 that males would receive approximately 65 fewer points than their female counterparts.

Reading scores are not statistically significantly impacted by the mother's educational background. The research suggests a mother's higher educational attainment (meaning higher than no education, primary or junior secondary education) is linked to roughly a 1.153 to 1.411 point rise in the reading score. The analysis indicates that, at a significance level of 1 per cent, a father's educational background has a positive association. There is a link between an additional ISCED level of education below the university level and an improvement in secondary school performance by 10.211 to 10.530 points. This result is logical because the parents can be motivated for their children to excel in secondary school.

At the one per cent significance level, computer use has a considerably favourable impact on reading test performance. In particular, there is a correlation between having access to a computer and an increase in reading test scores by roughly 30 points. These results are consistent with those obtained from the preceding dependent variables.

Additionally, having access to the internet affects secondary school performance. Secondary school performance rises from 43.626 to 45.139 as a result of internet access, at a 1 per cent significance level. Finally, at a one per cent significance level, the ESCS variable has a statistically significant effect of 44 points on reading test scores. The outcome makes logical sense because those pupils who have a high ESCS are provided with more resources and opportunities for learning.

After careful examination, it can be concluded that the Fixed Effects model can be a limited predictor of student secondary school performance using PISA data. This is due to its low explanatory power but mainly strong relationships. The model has limitations, such as the limited number of independent variables. Conclusions can be drawn from the model; however, further research should be conducted to comprehend the complexity of secondary school performance.

This study's hypothesis suggested that there would be a positive relationship between government funding and secondary school performance. However, the findings revealed a diminishing effect instead. Initially, increasing funding would improve secondary school performance. However, this effect slowly became negative. This would suggest that the relationship is more complex than it seems. Only increasing government funding to receive better test scores is inefficient.

5.2 Discussion

This paper revealed a significant, non-linear and diminishing effect of government funding on secondary school performance, providing an interesting perspective. According to the author's knowledge, the research on this topic has predominantly shown positive, negative, or no effect of government funding on secondary school performance, as evidenced by studies such as Mackenzie (2006) and Garen & Bray (1990). Due to the conflicting empirical evidence regarding the relationship, it is difficult to position the paper's results within the existing literature. Despite this difficulty, this study helps to reconcile these discrepancies by proposing a diminishing effect, contributing a new outlook. The results only partially align with the studies (Mackenzie (2006); Lubienski & Lubienski (2013)) that present only positive effects of government funding on school performance.

Instead, the study fits more into research conducted by Sisungo *et al.* (2014) that discusses the possibility of having an optimal level of government funding to positively help secondary school performance. This could potentially alleviate some of the responsibility of the government and redirect attention towards other critical factors influencing secondary school performance. Maybe it would be useful for researchers to conduct research on charter schools, which receive a variety of school funding, and perhaps future research could benefit from focusing on the magnitude of government funding. With variables such as ESCS strongly influencing secondary school performance, the effect of government funding is relatively minor, suggesting that addressing socioeconomic disparities may yield greater returns than increased government funding.

However, this study has limitations that need to be addressed. Firstly, the low R^2 of the FE models limits the accountability of the models. Secondly, since some countries and regions were unable to participate, the PISA data is not based on a random sample. Consequently, the findings may not be as generalisable and may, therefore, be applicable only to the OECD countries and regions that were included. Thirdly, a limitation of using school-level data is that the results are based on each school rather than on the individual students. Thus, the research cannot be as representative and applicable. Fourth, the study's generalisability is constrained by the lack of information about the adequacy of funding and the amount of funding the schools are receiving, unlike the study by Sisungo *et al.* (2014).

Additionally, the mother's education variable was consistently statistically insignificant. Upon careful consideration, this could be because the ISCED range may not have been broad enough as it does not include education levels beyond secondary education, such as Bachelor's or Master's degrees. This difference highlights a weakness in the dataset since it might not accurately reflect the variability and influence of mother education on secondary school performance. Therefore, this limitation may prevent the results from providing a complete picture of the relationship. Moreover, future research could extend the number of independent variables to see the full picture by considering, for example, peer effects, as they may significantly influence test scores (Ding & Lehrer (2007); Burke & Sass (2013)).

Other findings regarding the other independent variables that influence secondary school performance, including their statistical significance, are consistent with earlier research. For example, this study's findings regarding the male variable also illustrated that female students outperform male students in schools such as in Voyer & Voyer (2014). Moreover, students who had higher ESCS scores were associated with better secondary school performance. This emphasises how crucial it is to deal with the underlying causes of socioeconomic differences to be able to improve secondary school performance outcomes for all students regardless of their background. Future research could extend the number of variables by considering peer effects, for example, as they may significantly influence test scores.

Another research path to consider is conducting studies at the country level to estimate the effect of government funding on secondary school performance. This would allow for additional information about each country to be included to provide a more nuanced understanding. Secondly, researchers could explore the potential of combining different types (government and private) of school funding. Furthermore, the research suggests that there is a need for further exploration of the optimal percentage of government funding.

Chapter 6

Conclusion

This study investigates the influence of government funding in secondary schools on secondary school performance for the period 2006 to 2018. PISA test scores for mathematics, science, and reading were sourced from the standardised PISA assessments and used as the dependent variables. The explanatory variables included student and school characteristics, which were included in the three school-level datasets. Initially, the Ordinary Least Squares method was applied for the analysis. However, an alternative method had to be employed due to the panel nature of the data. The results of the Hausman test determined that the Fixed Effects method should be used. The Fixed Effects method was able to take unit-specific fixed effects and time-varying effects into account. To assess the effectiveness of different percentages of government funding, distinct model specifications were created for each dependent variable.

The comprehensive analysis revealed that there is only satisfactory evidence to imply that government funding has a non-linear relationship with secondary school performance. The low R^2 values indicate that there might be more explanatory variables that influence secondary school performance. The majority of the variables were statistically significant in multiple models and were associated with affecting the PISA test scores. The variables with the largest consistent impact across the models were the ESCS and internet variables. For example, a higher ESCS score was correlated with an approximate increase in the mathematics-dependent variable by 47 to 48 points at a 1 per cent significance level. Similarly, the internet variable was roughly linked with a 47 to 49-point increase in the mathematics score. Interestingly, the magnitude of the effect of the male variable on the reading test scores differed significantly from its impact on the science and math test scores.

The findings of this thesis are relevant for education policies as factors influencing secondary school performance are being analysed, which are important for future generations and policymakers. This thesis intends to offer a fresh new perspective to the existing research on the relationship between government funding and secondary school performance. This research opens up opportunities to investigate how combining various funding types can enhance secondary school performance. Lastly, this study suggests that government funding does positively improve secondary school performance, but only to an extent; therefore, it is crucial to carry out effective policies that cap a level or percentage of government funding to ensure that the schools efficiently utilise the funding.

Bibliography

ALEXANDER, K. L. & A. M. PALLAS (1983): "Private schools and public policy: New evidence on cognitive achievement in public and private schools." *Sociology of education* pp. 170–182.

ANDRABI, T., N. BAU, J. DAS, & A. I. KHWAJA (2010): "Are bad public schools public "bad?" test scores and civic values in public and private schools." *Bulletin of Education & Research* **24(2)**: pp. 45–56.

AZHAR, M., S. NADEEM, F. NAZ, F. PERVEEN, & A. SAMEEN (2014): "Impact of parental education and socio-economic status on academic achievements of university students." *European Journal of Psychological Research* .

BEDI, A. S. & A. GARG (2000): "The effectiveness of private versus public schools: The case of indonesia." *Journal of Development Economics* **61(2)**: pp. 463–494.

BENTLEY, B. & R. SIEBEN (2022): "Achievement in education–a study of senior high school achievement results based on gender." *International Journal of Research Publications IJRP* **110(1)**: pp. 317–335.

BURKE, M. A. & T. R. SASS (2013): "Classroom peer effects and student achievement." *Journal of Labor Economics* **31(1)**: pp. 51–82.

CHIATANTE, G., D. PELLITTERI-ROSA, E. TORRETTA, F. N. MARZANO, & A. MERIGGI (2021): "Indicators of biodiversity in an intensively cultivated and heavily human modified landscape." *Ecological Indicators* **130**: p. 108060.

COLEMAN, J., T. HOFFER, & S. KILGORE (1982): "Achievement and segregation in secondary schools: A further look at public and private school differences." *Sociology of education* **55(2)**: pp. 162–182.

Coleman, J. S. (1968): "Equality of educational opportunity." *Integrated education* **6(5)**: pp. 19–28.

Cornell, D., K. Shukla, & T. R. Konold (2016): "Authoritative school climate and student academic engagement, grades, and aspirations in middle and high schools." *Aera Open* **2(2)**: p. 2332858416633184.

Couch, J. F., W. F. Shughart, & A. L. Williams (1993): "Private school enrollment and public school performance." *Public Choice* **76(4)**: pp. 301–312.

Ding, W. & S. F. Lehrer (2007): "Do peers affect student achievement in china's secondary schools?" *The Review of Economics and Statistics* **89(2)**: pp. 300–312.

Esen, E. & O. C. Adigüzel (2023): "The effect of socioeconomic level on academic achievement: Evidence from pisa." *Journal of Theoretical Educational Science* **16(4)**: pp. 925–947.

Garen, J. E. & R. Bray (1990): "Educational test scores, education spending, and productivity in public education: National trends and evidence across states and over time, 1990–2015." *Education Spending, and Productivity in Public Education: National Trends and Evidence Across States and Over Time* **2015**.

Giambona, F. & M. Porcu (2018): "School size and students' achievement. empirical evidences from pisa survey data." *Socio-Economic Planning Sciences* **64**: pp. 66–77.

Goldhaber, D. D. (1996): "Public and private high schools: Is school choice an answer to the productivity problem?" *Economics of Education Review* **15(2)**: pp. 93–109.

Grace-Martin, M. & G. Gay (2001): "Web browsing, mobile computing and academic performance." *Journal of Educational Technology & Society* **4(3)**: pp. 95–107.

Grimes, P. W. (1994): "Public versus private secondary schools and the production of economic education." *The Journal of Economic Education* **25(1)**: pp. 17–30.

Gross, R. N. (2017): *Public vs. private: The early history of school choice in America.* Oxford University Press.

Hanushek, E. A. (1997): "Assessing the effects of school resources on student performance: An update." *Educational evaluation and policy analysis* **19(2)**: pp. 141–164.

Hanushek, E. A., S. Link, & L. Woessmann (2013): "Does school autonomy make sense everywhere? panel estimates from pisa." *Journal of Development Economics* **104**: pp. 212–232.

Haw, J. Y., R. B. King, & J. E. R. Trinidad (2021): "Need supportive teaching is associated with greater reading achievement: What the philippines can learn from pisa 2018." *International Journal of Educational Research* **110**: p. 101864.

Ibrahim, M., A. Norsuhaily, & I. Mudassir (2017): "Influence of parental education on academic performance of secondary school students in kuala terengganu." *International journal of academic research in Business and Social sciences* **7(8)**: pp. 2222–6990.

Idris, M., S. Hussain, & N. Ahmad (2020): "Relationship between parents' education and their children's academic achievement." *Journal of Arts & Social Sciences* **7(2)**: pp. 82–92.

Jimenez, E. & D. Cox (1989): "The relative effectiveness of private and public schools." *Technical report*, Journal of Development Economics.

Kapur, R. (2018): "Factors influencing the students academic performance in secondary schools in india." *University Of Delhi* **7(11)**: pp. 575–587.

Kushebayev, T. Z. & G. S. Nygymetov (2022): "The impact of state expenditure on education on student academic achievement." *Economics: the Strategy and practice* **17(4)**: pp. 201–211.

Li, Z. & Z. Qiu (2018): "How does family background affect childrenâ€™s educational achievement? evidence from contemporary china." *The Journal of Chinese Sociology* **5(1)**: pp. 1–21.

Lubienski, C. A. & S. T. Lubienski (2013): *The public school advantage: Why public schools outperform private schools.* University of Chicago Press.

Mackenzie, J. (2006): "Public school funding and performance." *University of Delaware* .

Martin, C. E. (2020): *An examination of school funding and ACT scores of high school juniors in Mississippi public school districts*. Belhaven University.

Morgan, J. & I. Sirageldin (1968): "A note on the quality dimension in education." *Journal of Political Economy* **76(5)**: pp. 1069–1077.

OECD (2018): "Pisa 2018 technical report." *Technical report*, Organisation for Economic Co-operation and Development (OECD).

Opatrny, M., T. Havranek, Z. Irsova, & M. Ščasnỳ (2023): "Class size and student achievement: A modern meta-analysis." *Technical report*, Institute of Economics Studies, Charles University.

Schleicher, A. (2017): "Seeing education through the prism of pisa." *European Journal of Education* **52(2)**: pp. 124–130.

Shero, J. A. & S. A. Hart (2022): "Methodological decisions and their impacts on the perceived relations between school funding and educational achievement." In "Frontiers in Education," volume 7, p. 1043471. Frontiers Media SA.

Sisungo, Z., L. Kaberia, & P. Buhere (2014): "An investigation on the relationship between the school level of funding and performance at kcse in mumias district, western province, kenya." *International journal of Educational Administration and policy studies* **6(2)**: pp. 15–22.

Sjøberg, S. & E. Jenkins (2022): "Pisa: a political project and a research agenda." *Studies in Science Education* **58(1)**: pp. 1–14.

Sokolakova, M. (2023): "The impact of european funding on learning outcomes of pupils in slovakia."

Suna, H. E., H. Tanberkan, B. Gür, M. Perc, & M. Özer (2020): "Socioeconomic status and school type as predictors of academic achievement." *Journal of Economy Culture and Society* **(61)**: pp. 41–64.

Teltemann, J. & R. Schunck (2016): "Education systems, school segregation, and second-generation immigrants' educational success: Evidence from a country-fixed effects approach using three waves of pisa." *International Journal of Comparative Sociology* **57(6)**: pp. 401–424.

ČEJKA, v. (2023): *Nechme mladé volit: Potenciální dopady snížení věku voličů na volby v Česku.* Bakalářská práce, Univerzita Karlova, Fakulta sociálních věd, Institut ekonomických studií.

Voyer, D. & S. D. Voyer (2014): "Gender differences in scholastic achievement: a meta-analysis." *Psychological bulletin* **140(4)**: p. 1174.

Wibowo, S. A. & M. I. Rukayah (2020): "Effect of public elementary school financial challenges on student achievement." *Journal Ilmu Pendidikan (JIP)* .

Williams, J. D. (2021): "Class size, funding and academic performance." *Journal of Education, Society and Behavioural Science* **34(1)**: pp. 66–78.

Wooldridge, J. M. (2010): *Econometric analysis of cross section and panel data.* MIT press.

Appendix A

Results

Table A.1: Random Effects Regression Results

Variable	**Mathematics**	**Science**	**Reading**
(Intercept)	437.15***	446.62***	468.84***
	(2.6898)	(2.6206)	(2.4785)
fund_gov	0.1058***	0.0836***	0.0031
	(0.0087)	(0.0084)	(0.0079)
stratio	-0.4853***	0.0282	-0.0746**
	(0.0574)	(0.0558)	(0.0246)
school_size	0.0245***	0.0217***	0.02198***
	(0.0007)	(0.0007)	(0.0006)
mother_educ	2.0739*	2.8966**	2.1637*
	(0.9535)	(0.9282)	(0.8903)
father_educ	-8.9867***	-12.047***	-14.357***
	(0.9162)	(0.8922)	(0.8596)
proportion_male	-13.429***	-26.034***	-63.080***
	(1.1224)	(1.0924)	(1.0688)
proportion_computer	30.121***	36.966***	47.571***
	(2.0304)	(1.9793)	(1.8856)
proportion_internet	27.984***	25.131***	22.050***
	(1.7496)	(1.7085)	(1.6191)
escs	47.448***	46.117***	45.322***
	(0.5079)	(0.4948)	(0.4802)
R^2	0.86106	0.87381	0.87354
Adj. R^2	0.86104	0.87379	0.87352
Chisq	43695.1	44798.2	55183.6
Num. obs	53532	53532	56924

Note: The full variable names are listed in Table A.5

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A.2: Hausman Test Results

Subject	χ^2 Value	p-value
Math	569.21	$< 2.2 \times 10^{-16}$
Science	755.41	$< 2.2 \times 10^{-16}$
Read	569.99	$< 2.2 \times 10^{-16}$

Table A.3: Variance Inflation Factor (VIF) Values for Each Variable

Variable	**Math**	**Science**	**Read**
Government Funding	1.053136	1.053136	1.052171
Stratio	1.105530	1.105530	1.105278
School size	1.098225	1.098225	1.098117
Mother education	1.261314	1.261314	1.178808
Father education	1.260380	1.260380	1.183875
Male	1.000743	1.000743	1.000744
Computer	1.707828	1.707828	1.707914
Internet	1.688336	1.688336	1.687686
Escs	1.530379	1.530379	1.462417

Table A.4: AIC (Akaike Information Criterion) Values for Each Variable

Model	**Math**	**Science**	**Read**
Government Funding	593136.6	591291.8	632764.0
Stratio	593386.2	591293.3	632782.7
School size	590285.5	588220.2	630047.4
Mother education	588258.2	585928.2	627406.1
Father Education	588900.5	586812.5	628388.9
Male	593094.4	583094.4	593094.4
Computer	572547.4	572547.4	572547.4
Internet	571868.7	571868.7	571868.7
ESCS	563388.1	560725.4	600584.1

Table A.5: Full Variable Names for Abbreviations

Full Variable Names for Abbreviations	**Variable**
fund_gov	Government Funding
stratio	Student-teacher ratio
school_size	School size
mother_educ	Mother's education
father_educ	Father's education
male	Male
computer	Computer
internet	Internet
escs	Escs