SCIENCES PO PARIS

MASTER THESIS

Evaluation of the impact of computer-aided instruction on student performance

---

Author:
Clémence LOBUT

Supervisor:
Dr. Denis FOUGÈRE

Master in Economics

May, 2017

Abstract

I study the impact of the use of ICT in schools on student's achievements in mathematics, reading and science, using data from PISA 2015 in 12 countries. PISA reports educational outcomes of 15-year-old students in OECD countries and partners. By OLS estimation, I find a negative impact of ICT use in school on student's achievement. However, this estimation might be biased due to endogeneity. By implementing propensity score matching to establish a causal relationship, I find that a high level of ICT use in school, compared to a low level, has small or non significant impact in some countries (Spain, France) but a large and negative one in others (Greece, Luxembourg, Poland)[1].

[1] I would like to thank Denis Fougère, my supervisor, for his helpful advices, as well as Guillaume Le Fur and Camille Renault for their insightful suggestions.

Contents

Chapter 1

Introduction

Being digitally literate is often described as a major asset to be an involved and enlightened citizen or in the labor market. This craze for new technologies reached education policy as many Ministries of Education expressed their willingness to open schools to the digital revolution. In Italy, the Ministry of Education launched the "National Plan for Digital Schools" (Piano Nazionale Scuola Digitale) to provide classrooms with ICT (Information and Communication Technologies) and foster new pedagogical practices (Avvisati *et al.*, 2013). In France, the "Digital Plan" (Plan numérique) for schools was launched in 2015, with the aim to give children "all keys to succeed in the digital era" [1] . This political will translated in an increase of the number of computers in classrooms, as observed by the OECD (Organisation for Economic Co-operation and Development). However, the organisation also points out that this raise remains moderate, as in 2012, 72% of 15-year-old students in OECD countries reported using computers at school (OECD, 2015).

Reaching the objective of a digital school requires investment in ICT equipment for classrooms, teacher training, etc. This implies important costs and resource allocation trade-offs. Hence, the issue of efficiency must be raised: does using computers and ICT at school has a positive impact on student achievements? Are benefits in terms of educational outcomes high enough to legitimize those investments? The aim of this analysis is to provide first keys of answer by measuring the impact of ICT use in schools on student performance estimates.

[1] French Ministry of Education website

To answer this question, I use 2015 PISA database. PISA ("Program for International Student Assessment") is an international study conducted by the OECD to measure 15-year-old students' performance in mathematics, reading and science. This database provides rich and high-quality data, allowing for comparisons across countries. As I do not have panel data or randomization in the access to ICT at school, I use Propensity Score Matching to evaluate the causal impact of computers at school on students' achievements.

In a subsample composed of 12 countries, I find a negative impact on students' achievements in all three subjects, especially in reading. However, those effects vary greatly between countries: they are not significant or very low in some (Spain, France) while much stronger in others (Greece, Luxembourg and Poland).

This analysis is constructed as follows. The first part explores results in the literature on this matter. In the second section, I present the data used and its specificities. The third part displays the results and limits of an OLS estimation of the relationship between ICT use at school and students' proficiency estimates. In the fourth section, I perform Propensity Score Matching to address the issues raised in the previous section and establish a causal impact of ICT in school and students' educational achievements. Finally, I perform robustness checks and discuss possible limits of the present analysis.

Chapter 2

Literature Review

This analysis focuses on computer-aided instruction (CAI), which designates the use of computer to teach math, reading, etc. Hence, I do not take into account teaching how to use computers and new technologies. CAI has been the focal point of the literature on ICT and education. This literature explores its impact on education with various scopes and at very different levels. If scientific research on this matter started in the 1970s, with the pioneering work of Supes in Stanford (1971), most of the research is quite recent, due to the digital revolution with the widespread use of ICT in the 90's.

In this concise literature review, I will only focus on studies that tried to assess the impact of the use of technologies on students' and pupils' school achievements. Many other impacts have been evaluated, such as attendance, creativity or interest in studying. Moreover, other researches have targeted special-need students, and are not analyzed here either.

One of the main motivations for research on this subject is the craze for new technology that generated important resource allocation for their implementation in classrooms and teacher training. The EU Commission reported that "the use of computers in European schools reached almost 100% saturation point in all member states" (Korte et al., 2006). As resource allocated for ICT in schools are often diverted from other traditional educational inputs, many raised the question of effectiveness and CAI's actual benefit on educational outcomes compared to other pedagogical solutions. As pointed out by Y. Zhao and J. Lei, "the hope that technology might bring significant improvement to education [. . .] has not been consistently supported by empirical evidence" (2009).

The OECD investigates this issue in their report, "Students, Computers and Learning" (2015), using data from PISA 2012. They find that the impact on student performance is mixed. If students receiving moderate levels of CAI have better education outcomes than those who did not receive any, they also have better educational outcome than those who use ICT very frequently, even when controlling for social background and students' demographic. Moreover, the use of computers at school have no impact on reducing the gap in achievements between students with advantaged background or disadvantaged background.

Overall results from the literature cannot be generalized, as no clear results stand out. Some studies find mixed outcomes, with no conclusive evidence of effectiveness of ICT use in school on students' achievements, as found by Condie and Munro (2007) who analyzed over 350 published sources. Others find small but positive effects, such as reported by Cheung and Slavin (2013) in their meta-analysis of 74 studies. They focus on studies with random assignment or matching, and find that CAI has modest but positive effect on schooling achievements compared to other traditional methods. On the other hand, Angrist and Lavy find small but negative effects in math in their study conducted in 2002.

This study is one of the most representative. Using a lottery organized by the Israeli State who provided funding for teacher training and hardware and software equipment, the authors set up a randomized experiment to assess the impact of CAI on pupil test scores in Math and Hebrew in 4th and 8th grade. From a simple OLS estimation with dummies for the levels of CAI intensity, they find no relationship except on 8th grade pupils in math, where the effect is negative. However, this effect disappears when they control for students' backgrounds. Then, they implement a 2SLS strategy where the instrument is a dummy indicating if the pupil is in a school that received funding before June 1996. They find a marginally statistically significant decline in test scores in 4th grade in math.

However, this study suffers from the same flaw as most of the studies in this literature, namely its low external validity, as conclusions are drawn from small-scale cases.

Fuchs and Woessmann (2004) study the impact of CAI at a much larger scale, using

PISA database. They analyze the relationship between students' achievements and the availability and use of computers at home and at school. They find a positive correlation between educational outcomes and availability of computers at home or at school. However, when controlling for family background and school characteristics, this relationship becomes negative for computers at home and insignificant for computers at school. When looking at the specific impact of different uses of computers at home, they find that there is a positive impact on students' achievements when pupils used ICT for educational and informative purposes. Moreover, they find that educational outcomes follow an inverted U-shaped curve depending on the intensity of computers' use at school. They give two interpretations: either this relationship suggests both a negative effect of CAI and an ability bias (low-achieving students do not have access to computers in schools), or the optimal level of CAI is low.

However, the main drawback of this study is that the interpretation is only based on conditional correlation, which cannot be interpreted as causal inference.

Important insights can be drawn from the study from De Witte and Rogge (2014), as they answer both the issue of scale and causal inference. They study the impact of ICT in schools on math using 2011 Trends in International Mathematics and Science Study (TIMSS) data. They construct control and treatment groups by Mahalanobis matching, based on students, teachers, school and regional characteristics. Their main result is that there is no significant effect and that not accounting for those characteristics may considerably bias the estimate of ICT's impact.

Hence, the main objective of this analysis is to provide a sound estimate of CAI's impact on education. Using PISA database and propensity-score matching, I address both the issue of scale and causal inference. Moreover, having estimates at country-level allows me to highlight variations in the impact of CAI on students' educational outcomes between countries.

Chapter 3

Data

1 PISA

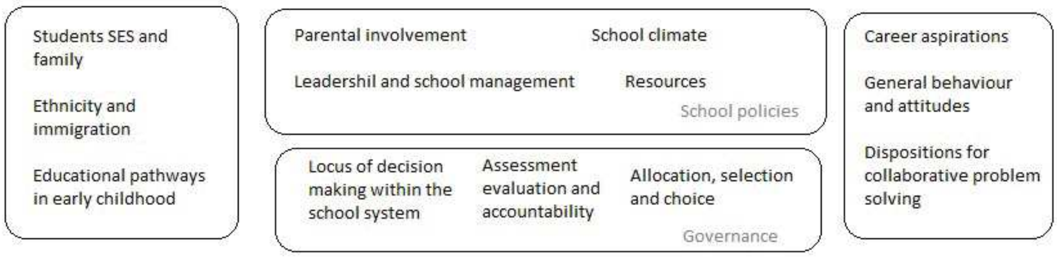
PISA is a worldwide study, implemented by the OECD. It aims to assess educational outcomes in mathematics, reading and science of 15-year-old students (between 15 years and 3 months old and 16 years and 2 months old). This age was chosen as it is close to the end of compulsory education, and school enrolment is almost universal at this level. In 2015, science was the major domain and collaborative problem solving was also tested.

The first PISA test was conducted in 2000 and has been repeated every 3 years since then. In 2015, 540,000 students were tested, from 72 countries and economies. They are representative of 29 million of 15-year-old students. Each country must draw a sample of at least 5,000 students. In small countries, such as Island and Luxembourg, when there are fewer students in this cohort, the entire age cohort was tested.

PISA's main objective is to answer the following question: "what is important for citizens to know and to be able to do?" The aim is not only to measure students' ability to reproduce what they have learnt, but more fundamentally to determine if they are able to extrapolate from that and to use their knowledge in unknown situations. Literacy in mathematics, reading and science is evaluated, over strict knowledge. Questions are based on real-world contexts and education's application to everyday-life problems and lifelong learning. For example, in reading, students are tested on their ability to "construct, extend and reflect on the meaning of what they have read across a wide range of continuous and non-continuous texts." (OECD, 2016).

Test items were designed in science, reading, mathematics and collective problem solving. They are a mixture of multiple-choice items and questions requiring students to construct their own answers. In total, they amounted to approximately 810 minutes of test. However, each student only answered part of those items for an amount of 2 hours approximately, different students taking different overlapping combinations of test items. Then, they answer a 35-minute-long background questionnaire on themselves, their family background, their school and learning experience. In 2015, 86.1% of the respondents took the test on computer and only 13.9% on paper. The assessment was conducted between March 1st, 2015 and August, 1st 2015.

One of the stated goals of PISA is to provide comparable data of education systems across OECD countries and partners, to enable countries to improve their education policies. PISA 2015 was specifically designed to combine "the existing approaches with new aspects of policy interest that currently guide the discussion on educational effectiveness and education policy decisions" (PISA 2015 Technical Report). To do so, 18 modules were defined: 5 on science-related topics and 13 on general topics as shown by the following graph:



Source: OECD (2016), "Modular structure of the PISA 2015 context assessment design"

PISA enables policy makers and researchers to identify the outcomes of various policy choices and to challenge deeply embedded academic processes. PISA reports have had strong impacts on national policies, as many states reformed their practices on assessment and curriculum standards and have incorporated PISA-like competencies in their educational program. For instance, Germany experienced what was called a

'PISA shock' in 2001, after the first results of the international test revealed that German pupils performed poorly compared to their peers in other developed countries. While education policies were for a long time decentralized, falling in the domain of the 16 Länder, all states agreed to create a national structure in charge of establishing a common national standard.

Nevertheless, some limits are to be acknowledged, especially in the uses and interpretations of PISA data. Indeed, in most OECD countries, PISA results are interpreted by the press, often not thoroughly enough, and actual education policies assessed in the lights of those results. Many studies have drawn unfounded interpretation of correlations. Education specialists have stressed that major policy decisions were drawn from flawed analysis, as too much extrapolation was made from PISA data. Therefore, it is important to stay careful when interpreting PISA results.

2 Data selection

Pisa databases from 2000 to 2015 are available to public on the OECD's website. They contain the full set of responses from individual students (background questionnaire and estimates of students' performance) and school principals, teachers (for 18 countries) and parents (for 22,5% of the students). In 2015, 519,334 students took the test. 72 countries participated, among which 35 are OECD countries and 36 partner countries and economies.

34,150 students from 12 countries were selected into the sample, based on four criteria. First, students from 26 countries did not take the ICT questionnaire, so data was missing about the use of ICT at school. Hence, 154,624 students from those 26 countries were withdrawn from the sample. Second, I constructed a binary treatment variable based on the continuous variable measuring the use of ICT at school. This treatment variable was equal to 1 for the 25% of the students who received the highest level of CAI, and 0 for the 25% who received the lowest. Students in the middle were discarded from the sample. Then, linked to the propensity score matching requirements, 20 countries failing the common support condition were also removed. The fourth selection round was much more arbitrary as it aimed to ease comparisons between

TABLE 3.1: Description of the 12 countries selected

Country	Observations	Percentage of final sample	OECD Country
Austria	2,648	7.27	X
Belgium	4,238	11.64	X
Bulgaria	2,703	7.42	
Spain	2,804	7.70	X
Finland	1,577	4.33	X
France	2,215	6.08	X
Greece	2,778	7.63	X
Israel	2,840	7.80	X
Italy	4,605	12.65	X
Luxembourg	2,201	6.04	X
Poland	2,238	6.15	X
Portugal	3,303	9.07	X
Total	34,150	100	11

countries, limiting them to Western Europe countries, or countries that were studied extensively in the education literature, such as Israel. Hence, the final sample was constituted of 34,150 students from 12 countries, described in the tables 3.1 and 3.2.

The ratio of treated and control units varies greatly between countries. Moreover, there are always more male students than female. The large majority of students are native from the country of the test, except for Luxembourg. Finally, the mean of the performance estimates are not the same as the ones reported by the OECD because the sample only contains a fraction of each country.

TABLE 3.2: Description of the sample (percentage)

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Lux.	Poland	Portugal	Total
Treatment status													
Treated	58,90	35,35	68,17	47,74	57,96	50,90	49,75	42,14	49,80	45,91	35,34	54,53	47,33
Control	41,10	64,65	31,83	52,26	42,04	49,10	50,25	57,86	50,20	54,09	64,66	45,47	52,67
Gender													
Female	43,68	46,48	42,57	46,36	36,74	47,19	44,61	45,64	46,45	48,17	45,57	43,77	45,96
Male	56,32	53,52	57,43	53,64	63,26	52,81	55,39	54,36	53,55	51,83	54,43	56,23	54,04
Immigration status													
Native	77,45	81,81	98,93	88,31	94,78	86,78	88,68	82,11	91,36	48,51	99,79	92,91	89,93
Second-Generation	13,49	9,70	0,55	1,93	2,41	9,28	7,47	13,21	3,18	30,56	0,17	3,03	5,23
First-Generation	9,07	8,50	0,52	9,76	2,82	3,94	3,85	4,68	5,46	20,93	0,05	4,05	4,84
Highest Education of parents (ISCED)													
None	0,61	0,67	0,11	1,07	0,56	0,49	0,13	1,23	0,08	2,42	0,00	4,86	0,63
ISCED 1	0,93	1,01	0,61	7,34	0,18	0,55	2,00	0,58	0,73	7,06	0,00	12,67	2,17
ISCED 2	2,49	2,98	3,80	17,04	1,27	7,02	5,88	2,07	18,47	8,37	6,38	24,40	10,65
ISCED 3B, C	23,78	4,84	7,55	6,94	10,69	14,46	3,79	6,82	4,85	5,21	17,85	6,40	10,36
ISCED 3A, ISCED 4	20,42	21,87	31,90	12,28	8,36	16,91	29,70	24,14	36,02	19,46	53,92	19,05	27,29
ISCED 5B	23,35	19,14	8,16	16,90	20,43	20,67	15,65	12,20	6,52	17,56	0,00	8,29	12,64
ISCED 5A, 6	28,41	49,49	47,87	38,43	58,52	39,90	42,85	52,96	33,34	39,92	21,86	24,33	36,25
Performance estimates (mean)													
Maths	487,23	504,91	435,80	481,53	491,09	483,40	446,85	467,59	483,40	481,27	500,48	483,51	483,58
Reading	471,92	495,37	426,44	487,31	498,08	487,82	460,53	474,57	487,82	472,94	497,87	486,81	483,92
Science	484,68	499,62	441,25	486,86	504,45	484,48	448,42	465,56	484,48	475,57	496,50	492,14	482,86

Many variables included in the dataset and used in this analysis were not directly drawn from students' responses to the questionnaire but were constructed from questionnaire items, as they cannot be observed directly (e.g. student's wealth or perceived emotional support for his or her parents). Those derived variables were constructed in three different ways:

- Simple indices were constructed from arithmetical transformation or recoding of questionnaire items

In this category, I use the variable on the highest education level of parents (*HISCED*) that is constructed from the items on mother highest education and father highest education. The classification is based on ISCED (International Standard Classification of Education) 1997, with 7 levels: (0) None, (1) ISCED 1 (primary education), (2) ISCED 2 (lower secondary), (3) ISCED Level 3B or 3C (vocational/prevocational upper secondary), (4) ISCED 3A (general upper secondary) and/or ISCED 4 (non-tertiary post-secondary), (5) ISCED 5B (vocational tertiary) and (6) ISCED 5A and/or ISCED 6 (theoretically oriented tertiary and post-graduate).

Data on the immigration background (*immig*) was also selected. This index was constructed from students' country of birth, their mother's and father's. Native students were born in the country of the test and have at least one parent born in the country, second generation students were born in the country of the test but their parent(s) were born abroad and first-generation students were born outside the country of the test and their parents were also born abroad. Students with missing responses for either the student or both parents were assigned missing values for this variable.

- Variables were derived from Item Response Theory (IRT):

  Six variables derived using IRT scaling are integrated in the dataset for this analysis:

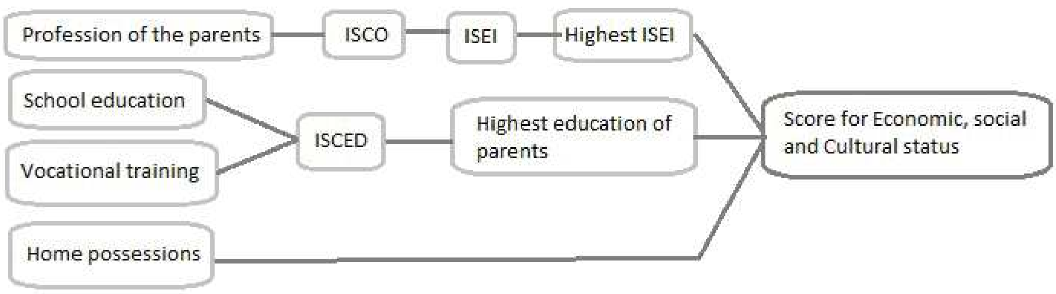
  - Wealth of the family (*wealth*), cultural possessions at home (*cultposs*), home education resources (*hedres*) and home possessions (*homepos*) are constructed from the reported availability of household items, including 3 country-specific items, seen as appropriate measures of the family wealth in the country context.

- Parental emotional support (*emosups*) if computed from items on whether parents are interested in school activities, support the students' educational efforts and achievements, support students when they are facing difficulties at school and encourage them to be confident.

- The use of ICT at school (*usesch*) is based on nine questions, with responses ranged from "never "never or hardly ever", "once or twice a month", "once or twice a week", "almost every day" to "every day": chatting online at school, using email at school, browsing the Internet for schoolwork, downloading, uploading or browsing material from the school's website (e.g. <intranet>), posting their work on the school's website, playing simulations at school, practicing and drilling, such as for foreign language learning or mathematics, doing homework on a school computer, using school computers for group work and communication with other students.

This variable *usesch* was the one used to construct the treatment variable

- Composite score for economic, social and cultural status



ISCO: International Standard Classification of Occupations; ISEI: occupation status of mother and father; HISEI: highest parental occupational status; ISCED: International Standard Classification of Education

Source: PISA 2015 technical guide

The score for Economic, Social and Cultural Status (ESCS) is a composite score built from the indicators of parental education, the highest parental occupation and home possession as a proxy for wealth via the Principal Component Analysis (PCA). Those elements were selected as socio-economic status is usually seen as based on education, occupational status and income.

3 Plausible values

In most international large-scale assessments, each student is not solicited on the full set of items but only on part of it, to reduce individuals' response time. Hence, different groups of students receive different but linked test booklets. This allows to assess a wide range of skill domains. However, it is not appropriate to use directly test scores (i.e. number of correct answers) as an estimate for student's performance because variations in test scores may come from differences in the difficulty of the test booklets, or because of some measurement error. To overcome this issue, plausible values were introduced for the first time for the 1983-1984 US National Assessment of Education Progress Data, based on Rubin's work (1987). Plausible values are "a representation of the range of abilities that a student might reasonably have" (Wu and Adams, 2002). They are generated by multiple imputations, using proficiency distribution and accounting for error or uncertainty at the individual level. Plausible values are drawn using an IRT model (1) and a latent regression model (2):

(1) The IRT model is used for item calibration, to provide information about test performance: regularities in the response pattern in a set of same-skill items can be used to characterized students and items on a common scale.

(2) Assuming that items parameters are fixed at the value obtained at step (1), a latent regression model is fitted to the data to get regression weights (Γ) and residual variance-covariance matrix (Σ). This latent regression model is based on a population model that uses test responses and answers to students' background questionnaire as covariates. Proficiency is not observed but is assumed to depend on the test item responses and on responses from the background questionnaire. This variable is treated as a missing variable following Rubin's approach (1987) and is used to derive plausible values.

Finally, ten plausible values are drawn for all students. They are random draws from the proficiency distribution using Γ and Σ, given the item responses and the background variables.

Chapter 4

Ordinary Least Squares

First, I estimate the effects of the treatment on proficiency estimates for mathematics, reading and science with a standard Ordinary Least Square (OLS) regression. I estimate the effect with no controls, and controlling for a large set of covariates. This set includes country fixed-effects to control for unobserved country-specific heterogeneity in the student proficiency. I also included dummies for gender, immigration status and highest level of education of the parents. Furthermore, the set of covariates contains indexes for the perceived parental emotional support, cultural possessions, home educational resources, home possessions and the economic, social and cultural status.

As proficiency estimates are given through plausible values, a specific procedure has to be followed, as described in PISA 2015 technical report. Let u be the number of plausible values. To compute β, the coefficient on the treatment and its variance, 4 steps must be followed. First, we have to compute β_u for each plausible value. Then, to compute the sampling variance of β for each plausible value, Var_u. The best estimate of β is the average of the ten β_u obtained from the ten plausible values. The best estimate of the variance of β is the sum of the average of the ten Var_u (uncertainty due to sampling from the population) and the variance among the β_u (uncertainty due to measurement error, as the true proficiency is not directly observed).

$$Var(\beta.) = \frac{\sum_{u=1}^{U} Var_u}{U} + (1 + \frac{1}{U})\frac{\sum_{u=1}^{U}(\beta_u - \beta.)^2}{U-1}$$

Results are displayed in the table 4.1. OLS results by country are in the annex. For the country fixed-effect, Luxembourg is the reference country , while for the highest education of the parents, the level ISCED 3A, 4 was excluded. The reference category for the immigration status is native students. Israel is not included as there were no observations for parental emotional support.

TABLE 4.1: OLS regression estimates on plausible values

	Mathematics		Reading		Science	
Constant	496.8***	519.9***	503.0***	502.5***	498.8***	517.1***
	(1.356)	(3.478)	(1.406)	(3.754)	(1.296)	(3.335)
Treated	-27.99***	-27.37***	-40.40***	-36.55***	-33.59***	-33.17***
	(1.923)	(1.891)	(2.238)	(2.086)	(1.937)	(1.917)
Country fixed effects						
Austria		3.540		-2.677		5.589
		(3.869)		(4.323)		(3.360)
Belgium		19.98***		18.46***		18.79***
		(3.387)		(3.745)		(3.204)
Bulgaria		-42.83***		-43.63***		-33.49***
		(4.557)		(5.267)		(4.637)
Spain		13.83***		27.63***		23.21***
		(3.538)		(3.875)		(3.348)
Finland		4.104		23.61***		22.63***
		(3.854)		(4.525)		(4.071)
France		9.820*		22.27***		15.28***
		(3.846)		(4.463)		(3.714)
Greece		-35.76***		-16.10***		-31.37***
		(3.823)		(4.360)		(3.993)
Italy		-6.184		-7.910		-11.91**
		(4.193)		(4.359)		(3.748)
Poland		16.89***		19.31***		14.82***
		(4.265)		(4.468)		(4.340)
Portugal		7.332*		17.75***		19.84***
		(3.433)		(3.910)		(3.400)
Gender						
Female		-16.22***		19.52***		-11.42***
		(1.951)		(1.907)		(1.772)

OLS regression estimates on plausible values (continued)

	Mathematics		Reading		Science	
Immigration status						
Second-generation		-15.48***		-13.75**		-19.54***
		(4.377)		(4.666)		(4.361)
First generation		-27.66***		-35.34***		-28.21***
		(3.999)		(4.462)		(3.995)
Parents highest education						
None		8.968		5.436		1.578
		(8.944)		(8.464)		(8.185)
ISCED 1		20.44**		20.28***		17.32**
		(6.497)		(5.819)		(5.501)
ISCED 2		1.319		0.0835		2.353
		(3.616)		(3.481)		(3.475)
ISCED 3B, C		-7.673*		-11.03**		-6.682*
		(3.241)		(3.371)		(3.348)
ISCED 5B		-10.87***		-9.519**		-9.513***
		(3.113)		(3.151)		(2.820)
ISCED 5A, 6		-19.87***		-22.26***		-20.10***
		(2.903)		(3.222)		(3.005)
Other						
Parental emo. support		-0.934		1.334		-0.341
		(0.803)		(0.945)		(0.802)
Cult. possessions		9.443***		15.80***		13.74***
		(1.262)		(1.187)		(1.109)
Home edu. ressources		5.605***		7.402***		6.114***
		(1.149)		(1.208)		(1.030)
Home possessions		-3.573		-10.60***		-8.155***
		(2.202)		(2.275)		(2.041)
Index of ESC status		36.58***		38.05***		37.63***
		(1.945)		(2.188)		(2.016)
Observations	34150	30032	34150	30032	34150	30032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Without controls, the OLS coefficient of the treatment displays a negative and significant impact of a high level of ICT use at school on students' educational outcomes in all three subjects. When adding controls, the absolute value of the coefficient on treatment decreases, so the negative effect of treatment on proficiency estimates is lower. Besides this negative impact, we can note that country fixed-effects are negative for Bulgaria and Greece, as well as for Finland, France and Portugal in reading and in science, while they are significant and positive for Belgium, Spain and Poland, but not statistically significant for Austria and Italy. Moreover, female students have on average lower proficiency estimates in math and science but higher in reading. As was expected, being from the second generation has a negative impact on proficiency estimates, and this impact is worse for first-generation students. More surprisingly, having parents with higher levels of education than ISCED 3B, 4 has a negative effect on achievements, while the effects of lower levels are not significant. However, this must be put in perspective with the large and positive impact of the economic, social and cultural status. Cultural possessions and home educational resources also have a positive impact, while the effect of home possessions is significantly negative in reading and science. Finally, parental emotional support has no significant impact on student achievement in our sample.

However, the coefficient on the treatment might be biased as there is a risk of endogeneity: the achievement estimate might be correlated with the residual. As Fuchs and Woessman state, "decision to use computers may not be random, but rather endogenously determined by students' ability. If our control variables do not fully control for student ability and if this ability is related to measured student performance, our estimates on computer use may well reflect this ability bias in addition to any causal effect of computer use" (2002). Indeed, schools with students having proficiency above the average might focus more on the use of ICT for education. Even if this issue is partly dealt with plausible values, the residual might still capture individual unobserved heterogeneity in terms of skills. Hence, the covariation between treatment and student's achievements may be driven by an unobserved factor that affects both. Therefore, randomized experiments or matching methods are required to measure the causal relationship between the use of ICT at school and student achievements.

Chapter 5

Propensity Score Matching

1 Theory

As data used here does not come from randomized trials, establishing a causal relationship requires to use instrumental variables or matching methods. No convincing instrument was found, so I used the latter. Rosenbaum and Rubin (1983) developed the propensity score matching (PSM) to establish an estimation of the treatment effect with observational data sets. As the probability of treatment assignment is not random but depends on other variables, using another method such as the OLS regression in the previous part provides biased estimates. PSM has become increasingly popular in medical trials and microeconomic policy evaluations, especially for the evaluation of training program or social welfare benefits. As we cannot observe simultaneously the potential outcomes of a given individual when receiving the treatment and when not receiving it, we cannot observe directly the causal effect.

The variable of interest is the average treatment effect of the treated (ATT). I also report the average treatment effect (ATE). However, as Heckman (1997) argued, its relevance is lower as it includes the effect of the treatment on persons for whom it was not intended. Hence, this estimate should be interpreted carefully, and only reflect the impact of the treatment if there were no spill-over effects such as peer effects. We can only identify the treatment effect when $E[Y(0)|D=1] - E[Y(0)|D=0] = 0$, where Y is the outcome and D the treatment status, i.e. if in absence of the treatment, outcomes for treatment and control groups would have been the same. It is only the case with randomized trials experiments as there is a risk of selection bias otherwise. Selection bias arises from differences between the treatment and control groups that affect the outcome, such as parents' wealth.

The aim of matching is to answer this issue of selection bias by matching treated and non-treated individuals with similar observable pre-treatment characteristics. We assume the unobservable characteristics play no role in the treatment assignment and outcome determination. To perform matching, several assumptions are required: unconfoundedness, the overlap condition and the conditional mean independence assumption. Unconfoundedness states that the assignment to treatment is independent of the outcome, conditional on the covariates X: $(Y(0); Y(1) \perp\!\!\!\perp D|X)$. It cannot be tested but must be plausible, based on the data and the institutional setup for the treatment implementation (Blundell et al., 2005). It implies that there is no omitted variable bias, once X is included in the regression.

Under unconfoundedness, we have:

$$E[Y(d)|X = x] = E[Y(d)|W = w; X = x] = E[Yobs|D = d; X = x]$$

Moreover, the overlap condition imposes that the probability of assignment is bounded by 0 and 1. Hence, for a given $X = x$, we can estimate both $E[Y(1)|X = x; D = 1]$ and $E[Y(0)|X = x; D = 0]$ and matching is feasible. Thus, we can estimate the ATE:

$$E[Y(1) - Y(0)|X = x] = E[Y(1)|X = x] - E[Y(0)|X = x]$$
$$E[Y(1)|X = x; D = 1] - E[Y(0)|X = x; D = 0] = E[Y^{obs}|X; D = 1] - E[Y^{obs}|X; D = 0]$$

Finally, the conditional mean independence assumption implies that the participation does not affect the mean of $Y(0)$: $E[Y(0)|X; D = 1] = E[Y(0)|X; D = 0] = E[Y(0)|X]$

Matching might bring less biased estimates than OLS for two reasons. First, because of the additional common support condition, it compares only comparable individuals. Moreover, matching is non-parametric and thus avoids misspecification of $E(Y(0)|X)$.

The balancing score b developed by Rosenbaum and Rubin (1983) is a function of observed covariates X such that the conditional distribution of X given b(X) is independent of assignment into treatment. The propensity score p is one of the possible balancing scores: it is the probability of receiving the treatment given observed characteristics X: $p(X) = Pr(D = 1|X = x) = E[D|X = x]$. If unconfoundedness and the balancing hypothesis hold, for any value of the propensity score, the difference between the treated and the untreated outcome is an unbiased estimate of the treatment effect at that value of the propensity score. Indeed, unconfoundedness allows that all bias due to selection effect can be removed by conditioning only on the propensity score. Moreover, because of the balancing hypothesis, two individuals with the same propensity score will have the same distribution of observable characteristics, independent of the treatment. Hence, treatment assignment is random at a given level of propensity score.

Propensity score can be estimated using standard probability model:

$$p(X_i) = Pr(D_i = 1|X_i) = F(h(X_i))$$

Where $h(X_i)$ is a function of covariates that can include linear and higher order terms such as interaction terms and F(.) is a cumulative distribution (such as probit). The covariates included in $h(X_i)$ should influence simultaneously the treatment and the outcome, and be unaffected by treatment. They should plausibly satisfy the unconfoundedness condition. Moreover, the inclusion of higher order terms is only determined by the need to pass the balancing property (Grilli and Rampichini, 2011).

Rubin and Thomas (1996) advice to include variables if there is any doubt on its impact on outcome. However, when including a new variable, there is a trade-off between the variance of the estimate propensity score and the plausibility of the unconfoundedness condition, as including more variables will increase both.

To assess the balancing property, the standard method is to test the independence of the treatment status and the covariates in each subsample with similar value of estimated propensity score. In each of those subsamples and for each of the covariates, the mean for the treated and for the control should be the same.

Once the propensity score is estimated, various matching algorithms can be used. Exact matching will match each treated unit with control units having the exact same propensity score. However, as $\hat{p}(x)$ is continuous, the probability to observe a treated individual and a non-treated individual with the exact same estimated propensity score is in principle zero. Hence, other algorithms are usually preferred, such as One-to-one Nearest-Neighbor matching or Kernel matching.

2 Application

The Propensity Score Matching is implemented following four steps, as suggested in Caliendo and Kopeinig (2008): (1) justification of unconfoundedness, (2) estimation of the propensity score, (3) selection of the matching algorithm and matching implementation, (4) verification of the balancing property.

1. Unconfoundedness

PISA database contains a high number of observations and high quality variables. As stated by Caliendo and Kopeinig (2008), those two elements allow to assume unconfoundedness.

2. Estimation of the propensity score

The propensity score is the probability of receiving the treatment given observable variables. As the treatment variable is binary, the propensity score is estimated using a probit regression.

$$p(X_i) = Pr(D_i = 1|X_i) = \phi(X'\beta)$$

Where $\phi(.)$ is the cumulative distribution function of the standard normal distribution. The propensity score was computed both using weights provided in the data and without weights, leading to very close results after matching.

The selection of the variables was based on three criteria, as defined by Rosenbaum and Rubin (1983). First, the variable should influence the decision of participation and the outcome variable. It should be unaffected by participation. Lastly, it should be unaffected by the anticipation of participation. Therefore, I only included covariates that fulfilled those three requirements and had only little missing data.

The list of variables included in the propensity score estimation varies slightly among countries, as the balancing condition could not be achieved using exactly the same set of covariates. The following set of variables was used: gender, immigration status, highest education of the parents (or only of the father or of the mother), number of school changes, parental emotional support, home educational resources, cultural possessions, home possessions, wealth and ESC status of the parents. Moreover, to pass the balancing test, I included higher order terms: the second-degree polynomial of the variable "home educational resources", based on its distribution, and interactions between the variables "wealth" and "home education resources" and "wealth" and "home possessions".

The χ^2 test was passed in the complete sample and in all subsamples. One of the advantages of PSM is that it highlights which covariates impact the probability to receive treatment, while this remains a black box in other methods such as diff-in-diff. To give an illustration, I selected four countries: Spain, France, Greece and Luxembourg. Within the complete sample, female students had a lower chance to have access to computers at school. This is also the case in all four countries under scope, although it is not statistically significant in Spain and in France. Moreover, for countries in which we had this information (here, only Spain), the higher the number of school changes, the greater the probability to receive the treatment. Moreover, having more home educational resources increases the chance to have access to ICT at school. The effect of home possessions and parental emotional support is ambiguous: the former has a positive impact within the complete sample but a negative one in Luxembourg and the latter has a positive impact in Spain but a negative one in Greece. Furthermore, having parents with a low level of education has a negative impact on the probability to have access to computers at school in Luxembourg. Finally, and interestingly, the probability to receive the treatment is greater with wealth while it decreases with the Economic, Social and Cultural status of the parents in France, Greece and Luxembourg.

For the outcome variable, I follow the analysis from De Witte and Rogge (2014) and use the average of the ten plausible values in each subject, as according to the Central Limit Theorem, their mean is a good approximation of the true value. However, matching was also performed using each singe plausible values, leading to the same overall conclusions [1].

To assess the balancing property, I used the command *pscore* in Stata. The property was satisfied for all subsamples, except for France and Belgium where only one variable was not balanced in one block. Tables 5.1 and 5.2 display means in the treatment and in the control group for the set of covariates selected, before and after matching, in Austria and Israel as examples. After matching, the balancing property is satisfied. Indeed, the results from the t-test show that there is no statistical difference between means of the two groups. Moreover, for all variables, the bias after matching is below 5% as usually recommended, except for interaction term between wealth and home educational resources in Israel but the bias is only of 5.6% and the difference is not statistically significant. Finally, the balancing property is usually assessed using two measurements introduced by Rubin (2001). Rubin's B is the standardized difference of means of the linear index of the propensity score in the treated and control group and should be below 0.25. Rubin's R is the ratio of treated to control variances of the propensity score and should be between 0.5 and 2.0. Here, Rubin's B and Rubin's R are respectively 0.058 and 1.14 for Austria and 0.094 and 0.93 for Israel, hence they satisfy the previous criteria.

3. Choosing a matching algorithm

Various matching algorithms can be used to match students from the control group to others in the treated group. They differ from each other in two ways: they define different neighborhoods for the treated individuals and assign different weights to control units matched with treated ones. A trade-off between bias and variance arises from this choice. However, when the sample grows, they should all return very similar estimates of the treatment effect.

Here, I use Epanechnikov Kernel Matching using the command *psmatch2* on Stata, where ATT is defined as:

[1] In two countries, results vary slightly from results using the average of the ten plausible values: in Spain, two plausible values in math returned positive but not significant difference after matching, as for France, three plausible values returned statistically significant difference in mathematics.

$$ATT = \frac{1}{N^T} \sum_{i \in (D=1)} [Y_i(1) - \sum_j w(i,j) Y_j(0)]$$

With the Epanechnikov Kernel mathing, w is defined as:

$$w(i,j) = \frac{K(\frac{p_j - p_i}{h_{nc}})}{\sum_{k \in (D=0)} K(\frac{p_k - p_i}{h_{nc}})}$$

And $K(u) = \frac{3}{4}(1 - u^2)\mathbb{1}(|u| < 1)$

Treatment group's outcomes are compared to those of the control group, with weight assigned to the latter depending on the distance between matched individuals in terms of propensity score. The bandwidth choice is quite important, as a trade-off between variance and bias arise (Silverman, 1986). Here, the bandwidth is 0.06.

One-to-one Nearest-neighbor and Mahalanobis matching were also performed using the command *teffects* on Stata and returned very close estimates, displayed in the annex.

TABLE 5.1: Balancing test - Austria

Variable	Unmatched Matched	Mean Treated	Control	%bias	%reduct |bias|	t-test t	p>|t|	V(C) / V(T)
Male	U	.3934	.49588	-20.7		-5.13	0.000	.
	M	.39368	.40014	-1.3	93.7	-0.35	0.728	.
Female	U	.6066	.50412	20.7		5.13	0.000	.
	M	.60632	.59986	1.3	93.7	0.35	0.728	.
Immigration status								
Native	U	.76095	.82768	-16.6		-4.07	0.000	.
	M	.76078	.76432	-0.9	94.7	-0.22	0.826	.
Second-generation	U	.15434	.10082	16.1		3.94	0.000	.
	M	.15445	.15199	0.7	95.4	0.18	0.857	.
First-generation	U	.08471	.07149	4.9		1.21	0.225	.
	M	.08477	.0837	0.4	91.9	0.10	0.919	.
Father Highest Education								
None	U	.01005	.01008	-0.0		-0.01	0.994	.
	M	.01006	.00899	1.1	-3224.7	0.29	0.771	.
ISCED 1	U	.0122	.00642	6.0		1.46	0.143	.
	M	.01221	.01155	0.7	88.5	0.16	0.871	.
ISCED 2	U	.04738	.05225	-2.2		-0.55	0.579	.
	M	.04741	.04534	1.0	57.3	0.26	0.794	.
ISCED 3B, C	U	.33453	.35564	-4.4		-1.10	0.272	.
	M	.33405	.35096	-3.6	19.9	-0.94	0.347	.
ISCED 3A, ISCED 4	U	.13496	.14299	-2.3		-0.57	0.566	.
	M	.13506	.13412	0.3	88.3	0.07	0.942	.
ISCED 5B	U	.25556	.18698	16.6		4.07	0.000	.
	M	.25575	.24846	1.8	89.4	0.44	0.658	.
ISCED 5A, 6	U	.20531	.24565	-9.7		-2.40	0.017	.
	M	.20546	.20059	1.2	87.9	0.32	0.750	.
Parental emo. support	U	.29431	.33349	-4.3		-1.07	0.283	0.97
	M	.29445	.28549	1.0	77.1	0.26	0.797	0.91
Cultural possessions	U	.09073	.1153	-2.3		-0.58	0.562	0.89
	M	.09111	.07038	2.0	15.7	0.52	0.600	0.93
Home edu. Resources	U	.17057	-.04588	24.5		6.06	0.000	1.02
	M	.16986	.16273	0.8	96.7	0.21	0.832	1.01
$Hedres^2$	U	.81682	.77315	4.1		1.01	0.312	1.19
	M	.81644	.80338	1.2	70.1	0.36	0.717	2.11
Hedres*wealth	U	.21041	.16746	5.4		1.33	0.182	1.19
	M	.20789	.17781	3.8	30.0	0.99	0.324	1.09
homepos*wealth	U	.65717	.51843	10.4		2.52	0.012	2.24
	M	.65428	.60827	3.5	66.8	0.85	0.398	1.46
Home possessions	U	.29297	.10652	22.2		5.49	0.000	1.10
	M	.29213	.27393	2.2	90.2	0.57	0.569	1.07
Wealth	U	.27737	.00302	33.4		8.26	0.000	1.04
	M	.27526	.25883	2.0	94.0	0.53	0.599	1.00
Status of ESC	U	.11407	.07397	4.7		1.17	0.244	0.93
	M	.11329	.09874	1.7	63.7	0.46	0.649	0.97

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B (%)	R
Unmatched	0.055	188.18	0.000	11.0	6.0	56.4	0.93
Matched	0.001	2.34	1.000	1.5	1.2	5.8	1.14

TABLE 5.2: Balancing test - Israel

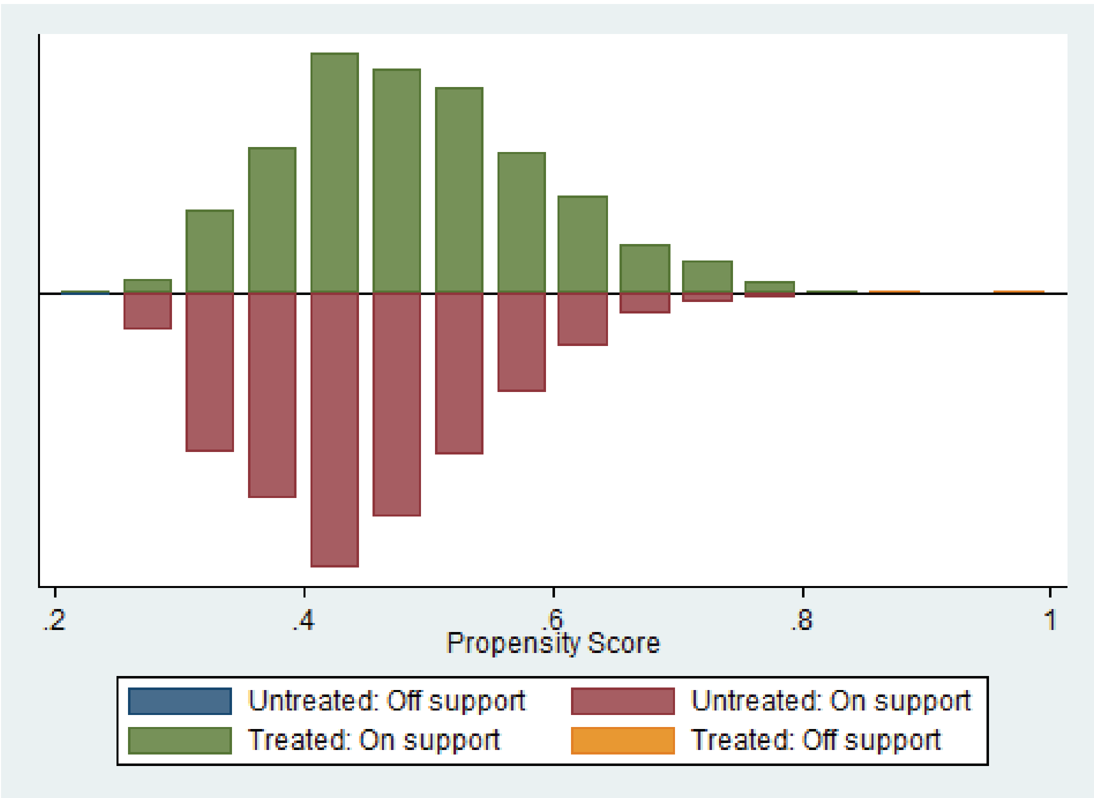
Variable	Unmatched Matched	Mean Treated	Control	%bias	%reduct |bias|	t-test t	p>|t|	V(C) / V(T)
Male	U	.46644	.53866	-14.5		-3.75	0.000	.
	M	.46644	.481	-2.9	79.8	-0.70	0.482	.
Female	U	.53356	.46134	14.5		3.75	0.000	.
	M	.53356	.519	2.9	79.8	0.70	0.482	.
Immigration status								
Native	U	.83133	.83143	-0.0		-0.01	0.994	.
	M	.83133	.82585	1.5	-5024.4	0.35	0.726	.
Second-generation	U	.11618	.13181	-4.7		-1.22	0.222	.
	M	.11618	.12033	-1.3	73.5	-0.31	0.757	.
First-generation	U	.0525	.03676	7.6		2.00	0.046	.
	M	.0525	.05383	-0.6	91.6	-0.14	0.886	.
Father Highest Education								
None	U	.00602	.01394	-8.0		-2.00	0.045	.
	M	.00602	.00563	0.4	95.0	0.12	0.901	.
ISCED 1	U	.0043	.00634	-2.8		-0.71	0.476	.
	M	.0043	.00405	0.3	87.6	0.09	0.925	.
ISCED 2	U	.02496	.01774	5.0		1.31	0.191	.
	M	.02496	.02395	0.7	86.0	0.16	0.875	.
ISCED 3B, C	U	.06282	.07668	-5.4		-1.40	0.162	.
	M	.06282	.06432	-0.6	89.2	-0.15	0.882	.
ISCED 3A, ISCED 4	U	.25731	.23954	4.1		1.07	0.287	.
	M	.25731	.2681	-2.5	39.3	-0.59	0.555	.
ISCED 5B	U	.13683	.11153	7.7		2.00	0.046	.
	M	.13683	.13304	1.2	85.0	0.27	0.789	.
ISCED 5A, 6	U	.50775	.53422	-5.3		-1.37	0.170	.
	M	.50775	.50091	1.4	74.2	0.33	0.742	.
Cultural possessions	U	.15731	-.01849	17.3		4.45	0.000	0.82
	M	.15731	.17145	-1.4	92.0	-0.33	0.741	0.77
Home edu. Resources	U	.35022	-.03916	41.0		10.54	0.000	0.88
	M	.35022	.36554	-1.6	96.1	-0.40	0.687	1.01
$Hedres^2$	U	.96631	.96423	0.2		0.05	0.963	0.57
	M	.96631	.96944	-0.3	-50.4	-0.08	0.938	0.98
Hedres*wealth	U	.3665	.31677	4.3		1.09	0.275	0.63
	M	.3665	.30144	5.6	-30.8	1.29	0.197	0.54
homepos*wealth	U	.95149	.77774	7.9		2.04	0.042	0.88
	M	.95149	.86045	4.1	47.6	0.89	0.372	0.60
Home possessions	U	.27854	-.01961	29.7		7.70	0.000	1.07
	M	.27854	.26154	1.7	94.3	0.40	0.686	1.03
Wealth	U	.26781	-.02309	30.4		7.89	0.000	1.09
	M	.26781	.22603	4.4	85.6	1.05	0.293	1.09
Status of ESC	U	.22756	.11717	13.0		3.35	0.001	
	M	.22756	.20625	2.5	80.7	0.62	0.537	

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B (%)	R
Unmatched	0.058	216.65	0.000	11.2	7.6	58.3	0.91
Matched	0.002	5.12	0.997	1.9	1.4	9.4	0.93

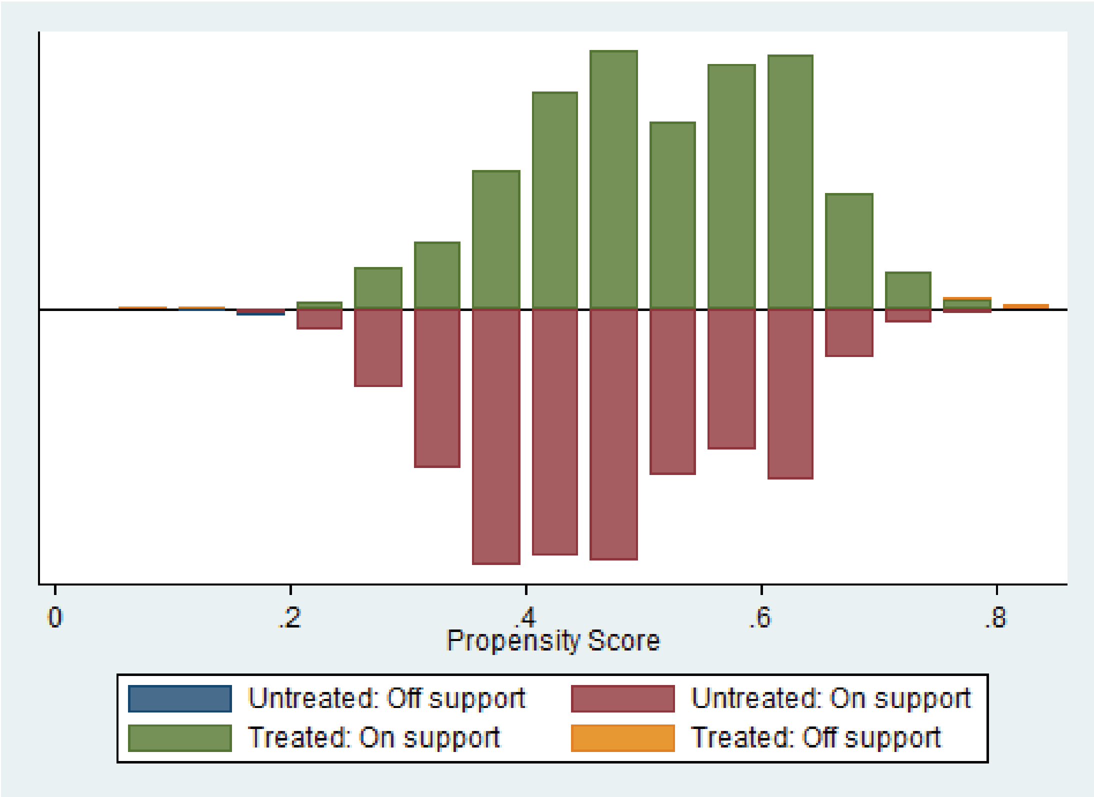
4. Common support condition

The common support condition is crucial as the average treatment effect of the treated is only defined on this region. I follow Lechner's recommendation (2008) and check for the common support by visualization of the propensity score distribution. From this visual analysis, 16 countries were discarded from the selection, as their common support was too small (the threshold chosen was at least 0.7 of common support).

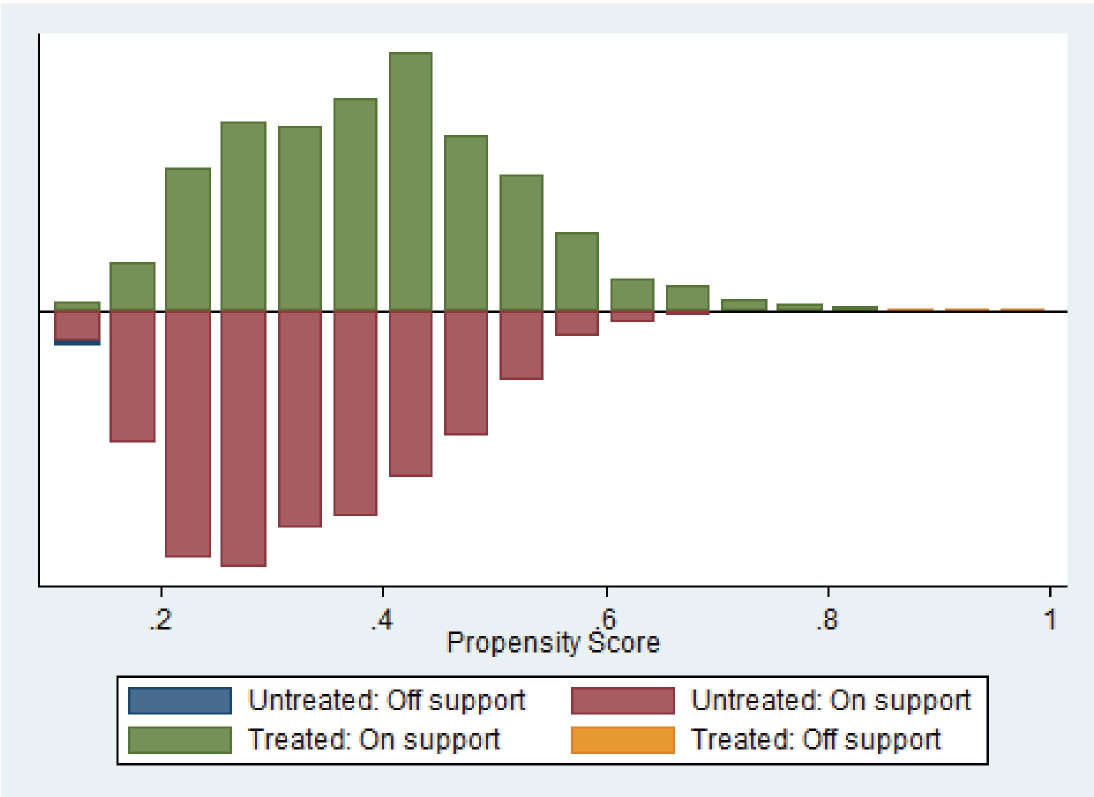
Following graphics show the distribution of the propensity score for the whole sample, as well as for 2 countries as examples (the one with the smaller common support, Spain, and the one with the larger common support, Poland). As individuals that fall outside of the common support region are discarded, their number should remain as small as possible, or the treatment effect estimate could be biased as it would only be computed from a sub-sample of the treated population that may not be representative. In this analysis, only a small number of students were discarded, representing 0.5% at worst (in the subsample for Belgium).



Distribution of the propensity score - Complete sample



Distribution of the propensity score - Spain



Distribution of the propensity score - Poland

3 Results

- Complete sample

Results from matching within the whole sample are displayed in table 5.3.

When matching using the whole sample, we find that the treatment, i.e. the use of ICT in schools still has a strong negative impact. However, this impact is somehow lower after matching. This is especially true in the case of reading where the ATT and the ATE in reading decrease in absolute value substantially by 6 points compared to the difference in the unmatched sample, when no weights are included, which represents a decrease of 11%.

However, as it can be seen from the OLS regression, the impact of the treatment differs greatly between countries. Therefore, I stratified the sample into subsamples by countries and applied the same matching strategy within each of those subsamples.

TABLE 5.3: Matching results - Complete sample

Subject	Weights	Sample	Treated	Controls	Difference	S.E.	T-stat
Maths	Weighted	Unmatched	460,94	501,28	-40,34	1,33	-30,29
		ATT	460,96	500,48	-39,52	1,38	-28,64
		ATE			-39,75	.	.
	Unweighted	Unmatched	460,94	501,28	-40,34	1,33	-30,29
		ATT	460,99	499,07	-38,08	1,38	-27,51
		ATE			-38,66	.	.
Reading	Weighted	Unmatched	453,93	507,97	-54,05	1,36	-39,76
		ATT	453,96	504,18	-50,22	1,41	-35,6
		ATE			-50,16	.	.
	Unweighted	Unmatched	453,93	507,97	-54,05	1,36	-39,76
		ATT	454,02	502,24	-48,21	1,42	-34,07
		ATE			-48,36	.	.
Sciences	Weighted	Unmatched	457,42	504,77	-47,35	1,36	-34,77
		ATT	457,45	503,64	-46,18	1,41	-32,74
		ATE			-46,34	.	.
	Unweighted	Unmatched	457,42	504,77	-47,35	1,36	-34,77
		ATT	457,50	501,99	-44,49	1,42	-31,43
		ATE			-44,96	.	.

Weighted Propensity score

Treatment status	Off support	On support	TOTAL
Untreated	1	8 150	8 151
Treated	5	7 187	7 192
Total	6	15 337	15 343

Unweighted Propensity score

Treatment status	Off support	On support	TOTAL
Untreated	3	8 148	8 151
Treated	6	7 186	7 192
Total	9	15 334	15 343

- By country

I report in the following tables the ATT and the ATE, where the propensity score was estimated using weighted probit regression. As with the whole sample, I used Kernel Matching. I do not report here the results for each subsample but selected four representative countries: Spain, France, Greece and Luxembourg. Results from the other countries are in the annex.

TABLE 5.4: Matching results - Spain

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Maths	Unmatched	492.88	490.97	1.92	3.05	0.63
	ATT	493.06	495.50	-2.44	3.21	-0.76
	ATE			-2.93		
Reading	Unmatched	493.69	501.86	-8.17	3.09	-2.65
	ATT	493.90	505.70	-11.80	3.24	-3.65
	ATE			-12.18		
Sciences	Unmatched	495.46	499.64	-4.18	3.22	-1.30
	ATT	495.651	503.80	-8.16	3.38	-2.41
	ATE			-8.72		

For Spain, the impact of the treatment becomes negative in math but is not significant, while the effect in reading is even more negative in the matched sample than in the unmatched one. We can observe the same trend in science, where the impact becomes significant at the 5% level. Estimates of the treatment effect in Austria followed the same pattern.

TABLE 5.5: Matching results - France

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Maths	Unmatched	486.88	498.57	-11.68	3.67	-3.18
	ATT	487.44	497.454	-10.02	3.91	-2.56
	ATE			-12.39		
Reading	Unmatched	488.93	508.22	-19.28	4.30	-4.48
	ATT	489.63	502.93	-13.30	4.57	-2.91
	ATE			-15.00		
Sciences	Unmatched	488.46	499.69	-11.23	4.04	-2.78
	ATT	488.88	496.31	-7.43	4.31	-1.72
	ATE			-10.01		

In France, matching brings smaller negative estimates of the treatment effect in all three subjects. This effect becomes not significant in science.

TABLE 5.6: Matching results - Greece

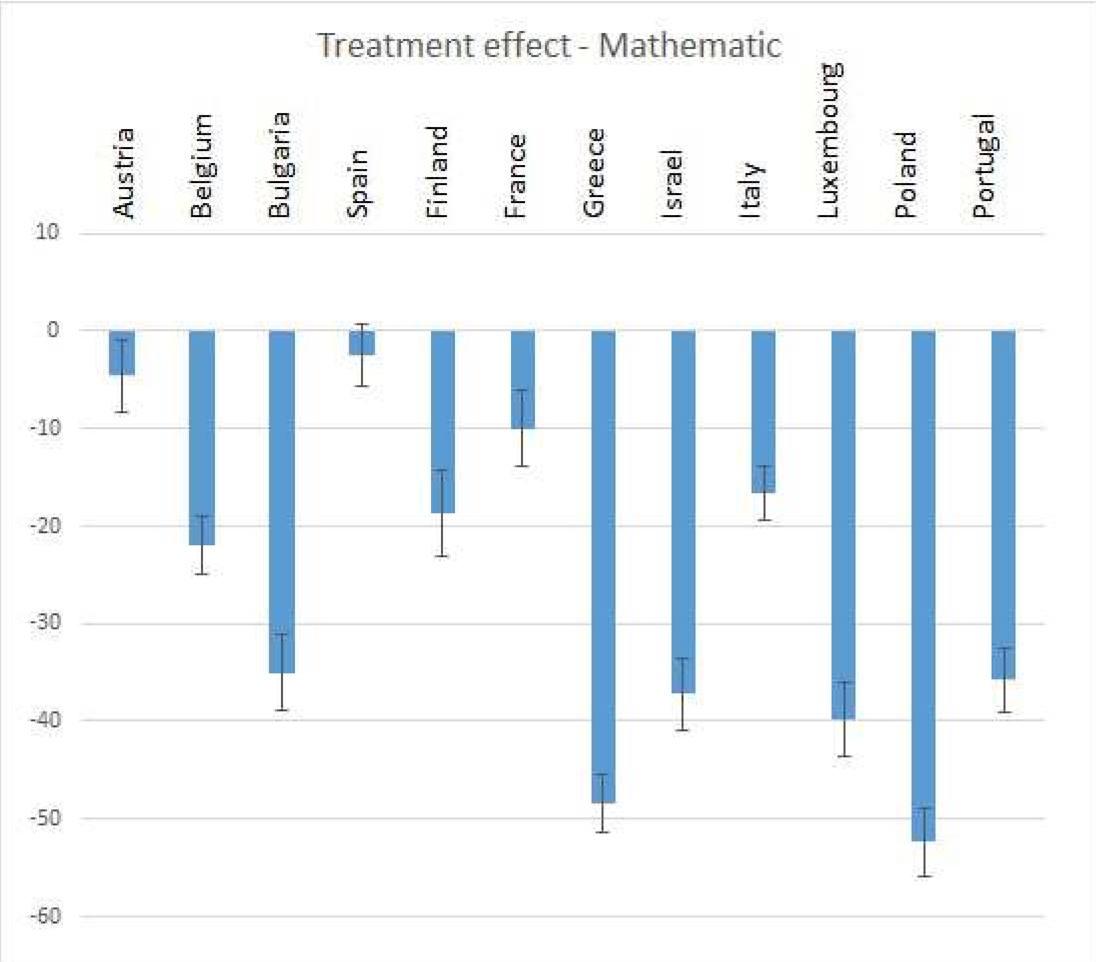
Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Maths	Unmatched	427.16	480.97	-53.80	2.82	-19.05
	ATT	427.18	475.57	-48.40	2.98	-16.25
	ATE			-48.80		
Reading	Unmatched	433.63	504.22	-70.59	2.91	-24.29
	ATT	433.74	492.49	-58.75	3.06	-19.18
	ATE			-59.30		
Sciences	Unmatched	423.18	488.64	-65.46	2.93	-22.33
	ATT	423.22	481.17	-57.95	3.09	-18.74
	ATE			-58.34		

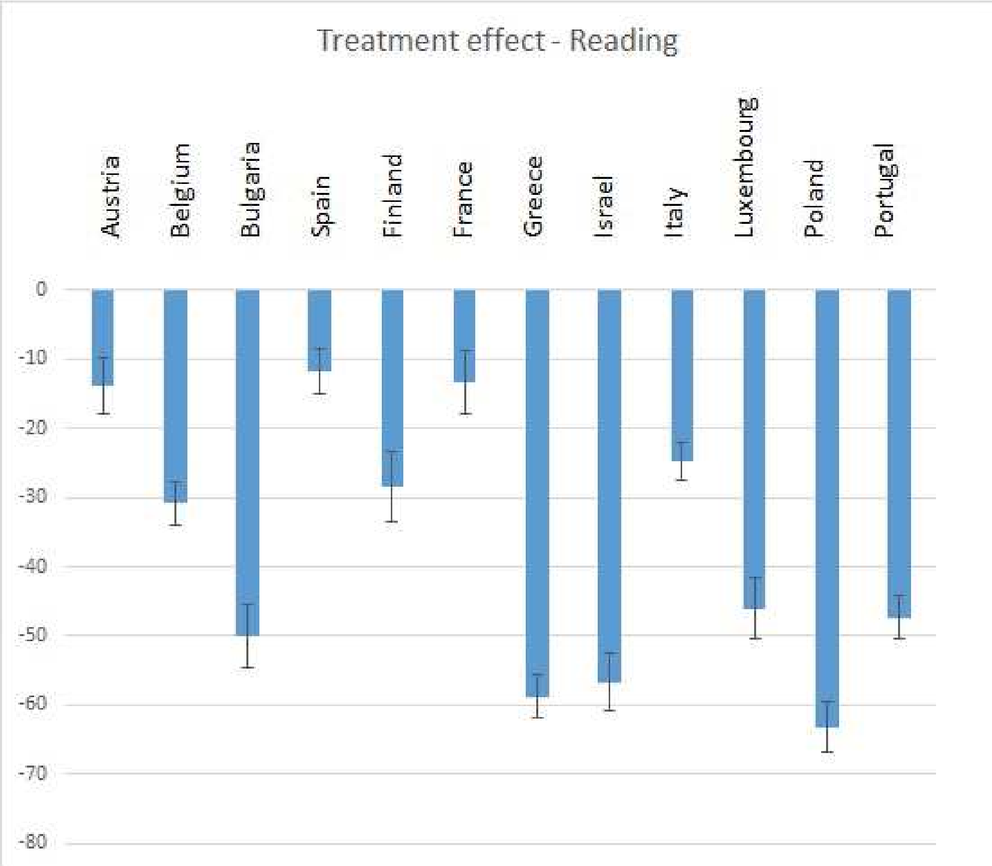
TABLE 5.7: Matching results - Luxembourg

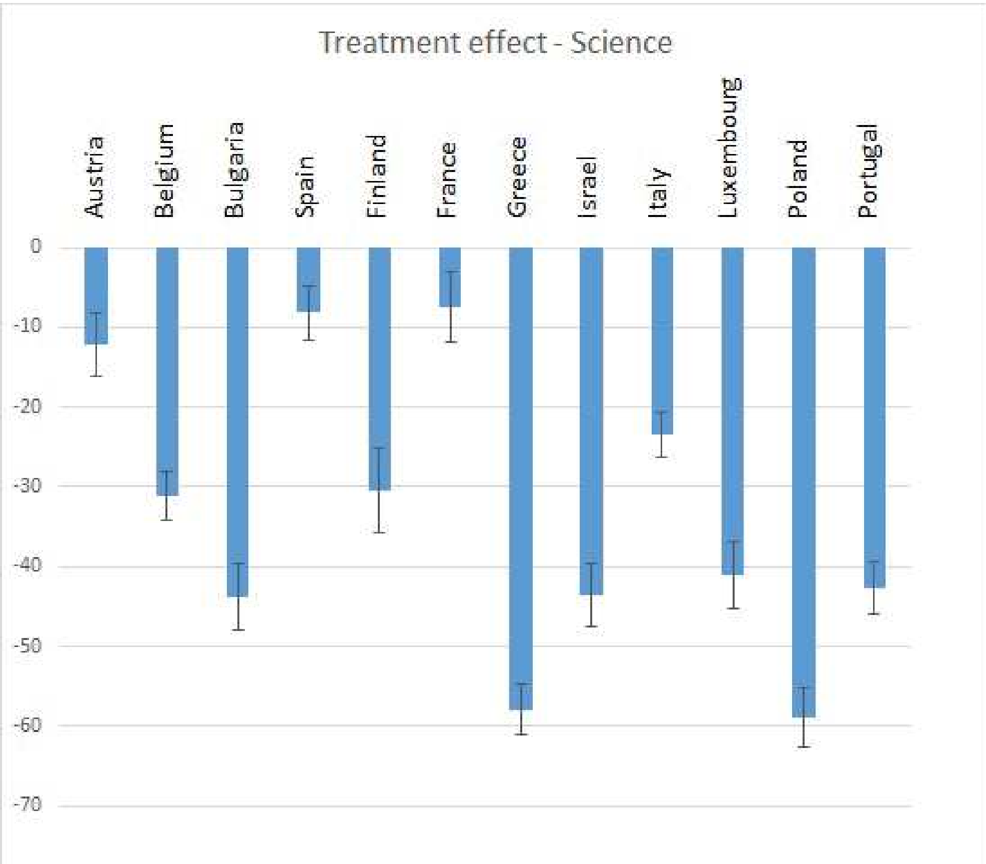
Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Maths	Unmatched	454.68	507.27	-52.59	3.56	-14.75
	ATT	454.68	494.52	-39.84	3.83	-10.39
	ATE			-42.50		
Reading	Unmatched	439.41	505.83	-66.42	4.13	-16.09
	ATT	439.41	485.44	-46.04	4.44	-10.36
	ATE			-48.92		
Sciences	Unmatched	446.69	504.04	-57.35	3.93	-14.58
	ATT	446.69	487.82	-41.13	4.23	-9.73
	ATE			-44.15		

On the other hand, Greece and Luxembourg's estimates are much higher in absolute value and are highly significant. The estimates change greatly from unmatched sample when PSM is implemented. Hence, if the effect of the use of ICT in schools is still greatly negative, this impact is not as strong as in the unmatched sample, especially in reading, where the estimate is reduced by 16,9% and 30,7%, respectively in Greece and in Luxembourg. The overall same trends can be observed in Israel, Poland and Portugal.

The following graphs give an overview of the results in all countries. If the impact of the use of ICT in school is quite homogeneous among subjects, we can observe that the negative effect on proficiency estimates in reading is larger in all countries.




Treatment effect - Reading
Austria
Belgium
Bulgaria
Spain
Finland
France
Greece
Israel
Italy
Luxembourg
Poland
Portugal
0
-10
-20
-30
-40
-50
-60
-70
-80


Treatment effect - Science
Austria
Belgium
Bulgaria
Spain
Finland
France
Greece
Israel
Italy
Luxembourg
Poland
Portugal
0
-10
-20
-30
-40
-50
-60
-70

Chapter 6

Robustness check and Discussion

1 Robustness check

I perform the same analysis but constructing the treatment variable differently. For the previous analysis, the control group was composed of the 25% bottom students in terms of ICT use in schools, and the treatment group of the 25% at the top. Here, I selected the 33% with the lowest level of ICT use in schools and 33% with the highest level. I conduct the same analysis with the same sample and with three subsamples (Spain, Poland and Finland). It leads to the overall same conclusions. However, the effect of CAI on students' performance is relatively smaller.

2 Discussion

Some issues could be raised to call into question the previous analysis. First, even if PISA data is of international renown for its richness and its quality, it comes from reported information. Students might have a different perception of the same reality and hence report it differently. Among two students who received the exact same amount of CAI, one might consider that he "never or hardly ever" uses ICT at school while the other will say that it happens "once or twice a month".

Moreover, the dataset is fairly large and complex. Even if the OECD publishes a great number of technical guides to help researchers applying statistical tools correctly, some difficulties were still encountered. The large number of variables available can be overwhelming and their selection (or non-selection) could be challenged. For example, I did not use data from teachers and schools, while they could have brought important insights, because that would have required to use three different and large datasets. Moreover, many other variables could have been used as treatment to answer different but interesting questions such as the use of ICT at home and its impact on student's achievement.

One of the main difficulties was linked to plausible values and more precisely their use in matching estimations. Further research should be conducted on this issue as the solution proposed here is not completely satisfactory. Nonetheless, using plausible values opens attractive research paths, as they reflect student's proficiency and not only his or her test scores. Hence, they may be seen as a measure of students' capabilities and thus answer part of the ability bias, at the center of research in education.

Chapter 7

Conclusion

Using data from PISA 2015 and implementing a propensity score matching strategy, I find a negative impact of the use of ICT in schools on student proficiency estimates in the selected sample. However, this effect varies greatly between countries under scope, as the impact is not statistically significant or very low in countries such as Spain or France, but much more important in other countries (Poland, Greece, Luxembourg). Hence, we could question the increasing craze for digital tools in education, and investments linked to it.

However, I only compared students receiving very low levels of CAI with students receiving very high levels of CAI. As pointed out in various studies (OECD, 2015; De Witte and Rogge, 2014), the impact of computers and ICT use in schools could follow a bell-shaped curve: CAI would only have a positive impact at intermediate level of intensity. Besides, computers and ICT might distract students from fundamental knowledge. Hence, the issue could be linked to teacher's training and how professors use digital tools.

Moreover, I only analyzed the impact on educational outcomes, omitting the importance of digital literacy nowadays. As underlined in the OECD report "Students, Computer and Learning", many of the potential benefits of ICT and computers are not measured by PISA. If using computers at school has a negative impact on school results, it may be the case that they have a positive impact in the labor market.

Finally, the impact of digital tools in schools on student may well depend on their habits at home. Interest in the digital divide has been growing in the recent years.

Indeed, if students do not have the same level of digital literacy due to an inequality of access or practices at home, CAI could have a negative impact on the perpetuation of inequalities at school.

Chapter 8

Appendix

Selection criteria

TABLE 8.1: Selection of the sample

Country	Nb of students	1	2	3	Selected	Nb. students selected
Albania	5215	X				
United Arab Emirates	14167	X				
Australia	14530		X			
Austria	7007				X	2 648
Belgium	9651				X	4 238
Bulgaria	5928				X	2 703
Brazil	23141			X		
Canada	20058	X				
Switzerland	5860		X			
Chile	7053		X			
Colombia	11795			X		
Costa Rica	6866			X		
Czech Republic	6894			X		
Germany	6504		X			
Denmark	7161		X			
Dominican Republic	4740	X				
Algeria	5519	X				
Spain	6736				X	2 804
Estonia	5587			X		
Finland	5882				X	1 577
France	6108				X	2 215

Sample selection (continued)

Country	Nb of students	1	2	3	Selected	Nb. students selected
United Kingdom	14157		X			
Georgia	5316	X				
Greece	5532				X	2 778
Hong Kong	5359		X			
Croatia	5809			X		
Hungary	5658			X		
Indonesia	6513	X				
Ireland	5741		X			
Iceland	3371			X		
Israel	6598				X	2 840
Italy	11583				X	4 605
Jordan	7267	X				
Japan	6647		X			
Korea	5581		X			
Kosovo	4826	X				
Lebanon	4546	X				
Lithuania	6525			X		
Luxembourg	5299				X	2 201
Latvia	4869			X		
Macao	4476			X		
Moldova	5325	X				
Mexico	7568			X		
Macedonia	5324	X				
Malta	3634	X				
Montenegro	5665	X				
Netherlands	5385		X			
Norway	5456	X				
New Zealand	4520		X			
Peru	6971		X			
Poland	4478				X	2 238
Portugal	7325				X	3 303
Argentina (BA)	1657	X				
Qatar	12083	X				

Sample selection (continued)

Country	Nb of students	1	2	3	Selected	Nb. students selected
B-S-J-G (China)	9841			X		
Massachusetts (USA)	1652	X				
Puerto Rico (USA)	1398	X				
North Carolina (USA)	1887	X				
Romania	4876	X				
Russian Federation	6036		X			
Singapore	6115		X			
Slovak Republic	6350		X			
Slovenia	6406			X		
Sweden	5458		X			
Chinese Taipei	7708		X			
Thailand	8249		X			
Trinidad and Tobago	4692	X				
Tunisia	5375	X				
Turkey	5895	X				
Uruguay	6062		X			
United States	5712	X				
Vietnam	5826	X				
Total country		26	20	14	12	
Total students	487004	154624	141447	108806	77649	34150

OLS results - Univariate

TABLE 8.2: Univariate OLS results - Mathematics

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Treated	-9.634	-21.85***	-35.33***	1.745	-10.15	-9.523	-55.96***	-50.54***	-30.89***	-52.88***	-58.52***	-46.14***
	(6.538)	(5.010)	(5.722)	(5.386)	(5.197)	(5.628)	(4.742)	(7.554)	(4.877)	(4.748)	(4.155)	(5.272)
Constant	492.9***	512.6***	459.9***	480.7***	497.0***	488.2***	474.7***	488.9***	498.7***	505.5***	521.2***	508.7***
	(4.254)	(3.616)	(5.580)	(3.891)	(4.408)	(4.123)	(3.801)	(6.341)	(3.754)	(3.129)	(3.188)	(3.940)
Observations	2648	4238	2703	2804	1577	2215	2778	2840	4605	2201	2238	3303

TABLE 8.3: Univariate OLS results - Reading

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Treated	-23.69***	-34.33***	-53.43***	-9.082	-19.43**	-17.19*	-73.63***	-71.87***	-44.84***	-67.25***	-76.46***	-58.47***
	(6.690)	(5.075)	(6.176)	(5.378)	(6.124)	(6.981)	(4.491)	(7.124)	(4.843)	(5.340)	(4.309)	(4.709)
Constant	485.9***	507.5***	462.9***	491.6***	509.3***	496.6***	497.2***	504.9***	498.4***	503.8***	524.9***	518.7***
	(4.468)	(3.595)	(6.070)	(3.459)	(5.129)	(4.266)	(3.691)	(6.167)	(3.531)	(3.247)	(3.364)	(3.617)
Observations	2648	4238	2703	2804	1577	2215	2778	2840	4605	2201	2238	3303

TABLE 8.4: Univariate OLS results - Science

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Treated	-18.17**	-30.73***	-46.52***	-4.893	-21.21***	-9.737	-67.42***	-55.78***	-40.18***	-58.03***	-67.26***	-54.16***
	(6.338)	(5.100)	(5.305)	(5.176)	(6.177)	(5.659)	(4.134)	(6.375)	(4.304)	(4.412)	(4.088)	(4.709)
Constant	495.4***	510.5***	473.0***	489.2***	516.7***	489.4***	482.0***	489.1***	495.3***	502.2***	520.3***	521.7***
	(3.930)	(3.572)	(5.644)	(3.376)	(4.633)	(3.815)	(3.480)	(5.497)	(3.152)	(2.972)	(3.293)	(3.646)
Observations	2648	4238	2703	2804	1577	2215	2778	2840	4605	2201	2238	3303

OLS results - Mathematics

TABLE 8.5: Regression table by country - Mathematics

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Treated	-13.31**	-24.16***	-38.15***	-5.068	-20.50***	-16.10**	-48.67***	-47.11***	-28.48***	-44.57***	-58.18***	-40.51***
	(5.134)	(4.205)	(5.378)	(4.847)	(5.043)	(5.031)	(4.676)	(6.479)	(4.384)	(4.611)	(4.054)	(4.881)
Female	-26.45***	-16.92***	-7.729	-16.79***	-2.592	-14.78**	-13.08**	-11.97	-25.09***	-18.01***	-16.35***	-17.29***
	(5.107)	(3.636)	(4.987)	(2.977)	(4.366)	(4.630)	(4.035)	(6.185)	(4.467)	(4.647)	(4.118)	(4.432)
Immigration status												
Second generation	-31.06***	-32.20***	-39.13	-11.64	-45.09**	-7.556	-13.85	12.47	-2.624	-5.050	-70.70	-7.719
	(8.052)	(5.946)	(23.45)	(11.43)	(14.17)	(8.164)	(8.647)	(7.010)	(13.17)	(5.168)	(60.60)	(12.07)
First generation	-59.57***	-33.46***	-39.87	-25.23***	-34.18*	-31.61**	-17.46	-31.56*	-17.01	-7.215	38.81	-21.20*
	(7.708)	(6.099)	(37.69)	(5.950)	(16.28)	(12.02)	(12.59)	(14.73)	(10.43)	(6.075)	(30.35)	(8.811)
Highest education of parents												
No education	-24.53	43.70*	0	-5.800	-24.08	6.583	22.56	49.43	32.46	96.39***	0	51.85***
	(34.33)	(17.64)	(.)	(18.35)	(40.67)	(22.85)	(58.74)	(25.39)	(36.98)	(15.66)	(.)	(15.76)
ISCED 1	-38.36	27.26	44.82	15.16	91.39*	25.13	-0.970	26.87	29.91	51.03***	0	36.32***
	(25.56)	(15.83)	(46.02)	(11.76)	(42.91)	(31.70)	(22.94)	(24.70)	(27.87)	(9.670)	(.)	(7.933)
ISCED 2	-21.91	10.72	-5.230	4.855	-8.579	-7.762	-10.91	-18.99	9.205	22.15**	-5.063	8.228
	(13.01)	(9.395)	(13.15)	(6.981)	(23.20)	(11.82)	(10.01)	(18.47)	(7.622)	(7.925)	(8.931)	(6.745)
ISCED 3B, C	-16.79**	-10.16	24.66***	-5.798	-0.518	-9.601	-6.276	-11.43	15.02	12.66	-12.53*	-9.947
	(6.313)	(6.852)	(7.407)	(8.672)	(9.873)	(7.395)	(10.55)	(6.583)	(9.681)	(9.906)	(5.499)	(9.173)
ISCED 5B	-33.31***	-11.72*	-13.02	-11.56	7.254	0.608	-19.94**	-13.96*	-69.90***	-33.85***	0	-19.01*
	(6.218)	(5.335)	(7.856)	(6.606)	(8.627)	(6.421)	(6.629)	(7.012)	(8.162)	(6.476)	(.)	(9.202)
ISCED 5A, 6	-25.64**	-30.66***	-23.87**	-16.29*	-1.483	-24.70***	-16.78*	7.758	-39.14***	-36.19***	6.880	-56.26***
	(8.024)	(5.631)	(7.279)	(7.698)	(9.167)	(6.672)	(6.533)	(6.929)	(6.638)	(7.428)	(6.802)	(8.483)

Regression table by country - Mathematics (continued)

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Parental emo. support	-3.226	-2.059	4.794*	-4.155	0.185	1.678	6.691**		-0.894	-0.884	-3.781*	-2.914
	(2.396)	(1.737)	(2.109)	(2.147)	(2.092)	(2.082)	(2.324)		(2.374)	(2.438)	(1.793)	(2.174)
Cult. possessions	7.708**	7.712***	12.70***	5.334*	8.033*	10.77***	8.393*	8.006*	8.724**	13.09***	13.41***	5.157
	(2.616)	(2.179)	(3.521)	(2.573)	(3.348)	(2.907)	(3.428)	(3.886)	(2.784)	(2.746)	(3.067)	(3.267)
Home edu. ressources	10.67***	7.501***	4.603	6.102*	1.275	4.592	9.620***	-0.154	6.897**	4.750	-1.422	2.849
	(2.835)	(2.178)	(2.964)	(2.509)	(3.583)	(3.263)	(2.860)	(3.046)	(2.302)	(2.582)	(2.520)	(3.348)
Home possessions	-4.721	-10.36**	-14.69**	0.959	-6.709	-2.469	-15.52***	-20.87***	-10.92*	-15.45***	2.906	-10.68*
	(4.937)	(3.723)	(4.712)	(4.514)	(6.204)	(5.058)	(4.457)	(6.129)	(5.159)	(4.152)	(5.036)	(5.351)
ESC status	33.54***	52.01***	48.00***	28.83***	32.84***	46.35***	32.57***	48.94***	47.49***	57.87***	21.86***	55.60***
	(6.432)	(4.003)	(5.427)	(4.601)	(5.502)	(5.331)	(4.793)	(6.493)	(5.162)	(5.385)	(4.812)	(5.183)
Constant	535.8***	543.0***	477.4***	517.9***	500.9***	524.6***	491.4***	486.4***	526.2***	526.6***	539.8***	546.9***
	(5.358)	(4.074)	(6.615)	(6.684)	(8.840)	(6.686)	(5.670)	(7.214)	(5.160)	(5.173)	(4.248)	(5.561)
Observations	2563	3903	2559	2705	1533	2124	2678	2740	4522	2062	2176	3207

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

OLS results - Reading

TABLE 8.6: Regression table by country - Reading

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Treated	-21.86***	-33.73***	-54.42***	-15.11**	-24.51***	-22.06***	-60.51***	-66.38***	-39.26***	-52.27***	-68.26***	-50.87***
	(5.887)	(4.254)	(5.632)	(4.901)	(5.205)	(6.518)	(4.133)	(6.296)	(4.261)	(5.016)	(3.916)	(4.093)
Female	17.92***	10.14**	33.87***	20.29***	34.22***	21.53***	18.91***	20.78**	9.037	11.01**	23.50***	6.305
	(5.400)	(3.799)	(4.951)	(3.845)	(5.141)	(5.127)	(4.272)	(6.425)	(5.317)	(4.116)	(3.880)	(3.869)
Immigration status												
Second generation	-23.00**	-33.24***	-43.22	-2.027	-34.63*	-9.194	-15.25	10.35	-5.896	-0.272	-86.14	15.64
	(8.590)	(6.662)	(30.19)	(14.18)	(14.78)	(9.853)	(9.778)	(7.573)	(12.38)	(5.084)	(51.77)	(11.23)
First generation	-63.99***	-23.86**	-52.54	-24.57***	-59.67**	-44.18**	-15.24	-44.60**	-46.60***	-8.872	18.56	-4.431
	(8.900)	(7.300)	(36.74)	(6.696)	(20.33)	(14.13)	(12.56)	(14.38)	(10.22)	(6.515)	(35.10)	(9.971)
Highest education of parents												
No education	-16.87	34.50	0	-25.92	-55.35	10.25	89.31*	63.00*	78.25*	113.6***	0	40.66***
	(30.29)	(23.93)	(.)	(16.22)	(44.10)	(28.36)	(43.25)	(29.81)	(39.55)	(20.14)	(.)	(11.44)
ISCED 1	-6.074	33.18*	79.16	8.636	119.8**	53.86	-3.452	8.260	12.83	59.83***	0	31.77***
	(32.77)	(14.41)	(44.34)	(10.97)	(37.48)	(41.19)	(20.09)	(26.28)	(25.28)	(12.45)	(.)	(8.983)
ISCED 2	-4.719	29.26**	-15.37	4.557	12.13	-9.690	-13.40	-18.50	5.978	14.18	-10.89	1.448
	(16.59)	(9.360)	(14.57)	(7.724)	(23.48)	(14.28)	(9.815)	(16.45)	(6.073)	(9.451)	(9.492)	(7.081)
ISCED 3B, C	-22.28**	-18.72*	9.579	-1.486	-7.750	-14.96	-25.64*	-13.51	14.76	24.85*	-16.63**	-9.109
	(7.151)	(7.874)	(8.857)	(9.357)	(11.61)	(8.031)	(10.46)	(8.110)	(10.18)	(11.30)	(5.636)	(10.06)
ISCED 5B	-40.93***	-12.13*	-5.594	-10.36	9.625	-2.831	-19.96*	-19.66**	-45.73***	-29.22***	0	-34.28***
	(7.193)	(5.416)	(9.420)	(6.525)	(10.48)	(7.556)	(7.873)	(7.174)	(9.168)	(7.629)	(.)	(10.02)
ISCED 5A, 6	-26.46**	-33.93***	-27.59***	-16.33*	-3.506	-31.68***	-17.90*	5.839	-35.93***	-40.77***	1.557	-59.48***
	(8.610)	(5.529)	(7.214)	(7.267)	(11.09)	(8.313)	(7.362)	(8.041)	(6.142)	(8.738)	(7.424)	(8.336)

Regression table by country - Reading (continued)

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Parental emo. support	-4.495	-1.487	10.04***	0.844	6.509**	3.320	7.065***		0.546	-3.845	-2.160	1.562
	(2.844)	(1.941)	(2.161)	(2.042)	(2.417)	(2.649)	(2.074)		(2.211)	(2.449)	(1.898)	(1.952)
Cult. possessions	15.22***	11.52***	19.25***	10.09***	11.65***	22.07***	9.418**	5.788	11.83***	17.12***	18.27***	11.47***
	(2.886)	(2.324)	(3.973)	(2.301)	(3.256)	(3.268)	(3.203)	(4.212)	(2.698)	(2.969)	(2.617)	(3.045)
Home edu. ressources	12.38***	11.14***	3.473	5.078*	-1.515	7.371*	12.00***	-1.543	9.240***	7.701**	-0.308	8.650**
	(2.726)	(2.191)	(3.214)	(2.397)	(4.004)	(3.426)	(2.323)	(3.554)	(2.587)	(2.672)	(2.625)	(3.092)
Home possessions	-17.31***	-16.26***	-19.78***	-8.416*	-10.09	-13.60*	-20.12***	-18.01**	-9.163	-22.27***	-4.388	-19.02***
	(5.043)	(4.263)	(5.075)	(4.224)	(6.351)	(6.236)	(3.773)	(6.794)	(5.467)	(4.463)	(5.045)	(4.744)
ESC status	41.68***	52.68***	55.83***	29.41***	31.28***	48.94***	37.09***	50.41***	43.47***	67.68***	25.60***	53.71***
	(6.843)	(4.004)	(5.280)	(4.492)	(5.613)	(6.755)	(4.751)	(6.842)	(4.898)	(6.191)	(4.616)	(5.030)
Constant	509.8***	526.8***	463.5***	511.8***	499.2***	521.9***	497.7***	489.3***	506.4***	508.6***	525.3***	546.1***
	(5.978)	(4.813)	(7.915)	(6.465)	(10.78)	(6.934)	(5.908)	(7.677)	(5.254)	(6.298)	(4.784)	(6.196)
Observations	2563	3903	2559	2705	1533	2124	2678	2740	4522	2062	2176	3207

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

OLS results - Science

TABLE 8.7: Regression table by country - Science

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Treated	-19.82***	-32.66***	-48.74***	-11.42*	-29.72***	-15.86**	-58.55***	-53.26***	-38.23***	-47.25***	-65.17***	-48.31***
	(5.226)	(4.265)	(4.805)	(4.453)	(5.578)	(5.291)	(3.946)	(5.654)	(3.883)	(4.282)	(3.770)	(4.276)
Female	-20.16***	-16.15***	3.396	-7.061*	6.034	-9.408*	-7.302*	-4.444	-24.80***	-16.51***	-12.08**	-18.01***
	(4.765)	(3.562)	(4.345)	(3.211)	(4.889)	(4.425)	(3.553)	(5.108)	(5.016)	(3.672)	(3.748)	(3.700)
Immigration status												
Second generation	-33.97***	-36.52***	-38.24	-13.56	-58.70***	-14.79	-16.83*	4.807	-4.015	-7.674	-66.69	-1.445
	(7.501)	(6.373)	(23.65)	(13.28)	(12.47)	(8.940)	(8.049)	(6.697)	(12.39)	(5.074)	(63.06)	(10.96)
First generation	-60.54***	-27.71***	-35.06	-23.03***	-35.36*	-43.16***	-15.86	-30.80*	-15.21	-6.579	13.02	-13.59
	(7.922)	(6.074)	(33.40)	(6.589)	(16.92)	(11.05)	(11.29)	(13.34)	(8.474)	(5.896)	(30.57)	(8.002)
Highest education of parents												
[1em] No education	-14.46	32.46	0	-25.10	-45.80	1.678	47.33	58.97*	58.60	104.8***	0	49.47***
	(27.75)	(23.44)	(.)	(15.42)	(36.69)	(21.79)	(45.83)	(24.20)	(33.34)	(17.88)	(.)	(10.29)
ISCED 1	-1.391	22.39	53.66	10.21	103.9***	48.21	9.095	26.03	11.72	56.79***	0	30.71***
	(24.39)	(14.85)	(37.98)	(10.41)	(26.93)	(30.12)	(19.26)	(24.37)	(21.50)	(11.41)	(.)	(6.734)
ISCED 2	-7.191	26.17**	-12.58	9.409	-1.609	-10.34	-8.208	-3.246	9.152	20.23*	-8.100	4.246
	(15.60)	(9.304)	(11.58)	(7.511)	(22.77)	(12.25)	(9.238)	(15.31)	(6.004)	(8.633)	(8.520)	(5.563)
ISCED 3B, C	-22.42***	-10.92	11.15	2.883	-6.347	-7.891	-13.61	-4.269	13.02	21.81*	-12.61*	-4.165
	(5.905)	(7.800)	(6.761)	(8.765)	(10.36)	(7.311)	(10.11)	(7.519)	(8.404)	(10.00)	(5.285)	(8.134)
ISCED 5B	-39.54***	-10.85	-14.64	-11.39	8.011	-2.044	-14.88*	-7.989	-55.06***	-33.63***	0	-22.35**
	(5.873)	(5.613)	(7.814)	(6.406)	(8.748)	(6.168)	(6.369)	(7.485)	(8.815)	(7.022)	(.)	(7.969)
ISCED 5A, 6	-26.50***	-30.32***	-27.82***	-20.92**	-10.43	-25.60***	-14.43*	11.82	-40.00***	-43.47***	8.871	-53.06***
	(7.943)	(5.558)	(6.546)	(6.748)	(8.803)	(7.137)	(6.084)	(7.065)	(6.292)	(7.703)	(7.173)	(7.672)

Regression table by country - Science (continued)

	Austria	Belgium	Bulgaria	Spain	Finland	France	Greece	Israel	Italy	Luxembourg	Poland	Portugal
Parental emo. support	-3.330	-2.360	5.990**	-2.441	3.995	1.192	6.989***		-0.871	-3.215	-3.098	1.005
	(2.472)	(1.718)	(2.180)	(2.043)	(2.231)	(2.065)	(1.875)		(2.110)	(1.911)	(1.965)	(1.906)
Cult. possessions	10.13***	9.883***	16.01***	9.417***	12.63***	14.97***	8.854**	9.676**	13.98***	17.64***	19.35***	8.616**
	(2.693)	(2.233)	(3.529)	(2.190)	(3.473)	(2.891)	(3.191)	(3.714)	(2.425)	(2.710)	(2.943)	(2.699)
Home edu. ressources	11.05***	8.487***	3.082	4.834*	-0.650	5.492	10.95***	0.181	10.21***	4.665	-3.049	4.779
	(2.684)	(2.084)	(2.808)	(2.335)	(3.675)	(2.923)	(2.208)	(2.975)	(2.011)	(2.395)	(2.584)	(3.137)
Home possessions	-12.92**	-11.26**	-15.98***	-3.507	-12.71	-7.355	-17.42***	-18.11**	-17.56***	-21.56***	-1.594	-15.84***
	(4.787)	(4.198)	(4.178)	(3.889)	(6.654)	(5.105)	(3.927)	(5.888)	(4.807)	(4.356)	(5.314)	(4.537)
ESC status	39.99***	53.47***	50.48***	31.61***	37.87***	47.95***	34.28***	47.14***	46.06***	66.91***	23.66***	53.40***
	(6.048)	(3.960)	(4.490)	(4.177)	(5.817)	(5.954)	(4.225)	(6.132)	(5.000)	(5.770)	(4.697)	(4.697)
Constant	538.1***	539.8***	488.9***	524.0***	522.2***	525.9***	493.5***	480.6***	520.2***	525.4***	536.7***	557.8***
	(5.095)	(4.516)	(6.718)	(5.773)	(8.647)	(6.357)	(4.686)	(5.881)	(4.806)	(5.745)	(3.966)	(4.959)
Observations	2563	3903	2559	2705	1533	2124	2678	2740	4522	2062	2176	3207

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

PSM with Nearest-neighbor and Malahanobis matching

TABLE 8.8: Matching results - Nearest Neighbor Matching

		Coef.	AI Robust Std. Err.	z	P>|z|	[95%	Conf. Interval]
	Mathematics						
ATE	treated						
	(1 vs 0)	-37.41671	1.460543	-25.62	0.000	-40.27932	-34.5541
	Reading						
ATE	treated						
	(1 vs 0)	-46.97515	1.492006	-31.48	0.000	-49.89943	-44.05087
	Science						
ATE	treated						
	(1 vs 0)	-44.10004	1.511149	-29.18	0.000	-47.06184	-41.13824

TABLE 8.9: Matching results - Mahalanobis Matching

		Coef.	AI Robust Std. Err.	z	P>|z|	[95%	Conf. Interval]
	Mathematics						
ATE	treated						
	(1 vs 0)	-37.56765	1.415139	-26.55	0.000	-40.34127	-34.79403
	Reading						
ATE	treated						
	(1 vs 0)	-46.82108	1.428231	-32.78	0.000	-49.62036	-44.0218
	Science						
ATE	treated						
	(1 vs 0)	-43.9557	1.455705	-30.20	0.000	-46.80883	-41.10257

Kernel matching results - Other countries

TABLE 8.10: Matching results - Austria

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	486.554567	494.462993	-7.90842649	3.45820676	-2.29
	ATT	486.545499	491.140347	-4.59484746	3.68768741	-1.25
	ATU	494.415728	485.647222	-8.76850621	.	.
	ATE			-6.42681617	.	.
Reading	Unmatched	464.222182	487.098946	-22.8767638	3.72995708	-6.13
	ATT	464.182797	477.971478	-13.7886811	3.99897648	-3.45
	ATU	486.945306	469.894111	-17.0511948	.	.
	ATE		-15.2207155	.	.	
Science	Unmatched	479.110729	497.039633	-17.9289035	3.64457851	-4.92
	ATT	479.078939	491.200692	-12.1217529	3.91654917	-3.10
	ATU	496.937124	480.629674	-16.3074505	.	.
	ATE			-13.9590059	.	.

TABLE 8.11: Matching results - Belgium

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	499.391552	522.14734	-22.7557876	2.78631299	-8.17
	ATT	499.427574	521.323488	-21.8959146	2.96643203	-7.38
	ATU	521.917818	494.629687	-27.2881313	.	.
	ATE			-25.3126961	.	.
Reading	Unmatched	481.819572	517.374742	-35.55517	2.85316487	-12.46
	ATT	481.896222	512.77498	-30.8787579	3.04887501	-10.13
	ATU	517.078975	480.567975	-36.5110003	.	.
	ATE			-34.4476318	.	.
Science	Unmatched	487.853003	520.1945	-32.3414965	2.95209991	-10.96
	ATT	487.91591	519.084102	-31.1681922	3.13535901	-9.94
	ATU	519.979538	484.127757	-35.851781	.	.
	ATE			-34.1359512	.	.

TABLE 8.12: Matching results - Bulgaria

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	428.685635	466.591118	-37.9054835	3.5328747	-10.73
	ATT	428.881512	463.871639	-34.9901273	3.96297537	-8.83
	ATU	466.711258	430.251001	-36.4602566	.	.
	ATE			-35.4578957	.	.
Reading	Unmatched	415.638222	472.532212	-56.8939898	4.04740131	-14.06
	ATT	415.978375	466.00861	-50.0302345	4.49257347	-11.14
	ATU	472.65536	420.081045	-52.5743155	.	.
	ATE			-50.8397148	.	.
Science	Unmatched	431.64752	481.049918	-49.4023974	3.69104603	-13.38
	ATT	431.902258	475.666982	-43.7647243	4.19531474	-10.43
	ATU	481.188016	434.211609	-46.9764072	.	.
	ATE			-44.7866234	.	.

TABLE 8.13: Matching results - Finland

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	488.445565	501.157346	-12.7117808	3.82137998	-3.33
	ATT	488.766479	507.468114	-18.7016352	4.37810919	-4.27
	ATU	501.138158	481.797521	-19.3406364	.	.
	ATE			-18.9676768	.	.
Reading	Unmatched	492.749538	515.46755	-22.7180114	4.43269904	-5.13
	ATT	492.954864	521.36368	-28.4088156	5.0305425	-5.65
	ATU	515.510902	488.607949	-26.9029531	.	.
	ATE			-27.781865	.	.
Science	Unmatched	498.123837	522.189264	-24.0654268	4.63514302	-5.19
	ATT	498.34879	528.856638	-30.5078482	5.30749018	-5.75
	ATU	522.188864	492.488647	-29.7002164	.	.
	ATE			-30.1715989	.	.

TABLE 8.14: Matching results - Israel

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	443.434934	489.980512	-46.5455782	3.53334474	-13.17
	ATT	443.434934	480.656945	-37.2220111	3.76739498	-9.88
	ATU	490.281325	443.912102	-46.3692232	.	.
	ATE			-42.4700814	.	.
Reading	Unmatched	440.287271	508.09455	-67.8072783	3.80536947	-17.82
	ATT	440.287271	496.954865	-56.6675937	4.06664745	-13.93
	ATU	508.591928	442.877083	-65.714844	.	.
	ATE			-61.8583125	.	.
Science	Unmatched	439.463652	491.567643	-52.1039907	3.66796081	-14.21
	ATT	439.463652	483.053229	-43.589577	3.90814349	-11.15
	ATU	491.885267	439.572982	-52.3122849	.	.
	ATE			-48.5940947	.	.

TABLE 8.15: Matching results - Italy

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	484.387036	506.733056	-22.3460201	2.59744319	-8.60
	ATT	484.387036	500.924464	-16.5374279	2.78351808	-5.94
	ATU	506.742854	489.481377	-17.2614773	.	.
	ATE			-16.8844575	.	.
Reading	Unmatched	470.242236	505.239735	-34.9974989	2.55318022	-13.71
	ATT	470.242236	495.036401	-24.7941649	2.72302237	-9.11
	ATU	505.211562	479.020715	-26.1908463	.	.
	ATE			-25.4635803	.	.
Science	Unmatched	473.325775	504.399774	-31.0739991	2.60470959	-11.93
	ATT	473.325775	496.732351	-23.4065757	2.79003978	-8.39
	ATU	504.362839	480.06711	-24.2957285	.	.
	ATE			-23.8327377	.	.

TABLE 8.16: Matching results - Poland

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	464.782925	523.055311	-58.2723854	3.4132296	-17.07
	ATT	464.995242	517.37518	-52.379938	3.52941272	-14.84
	ATU	522.682654	465.279201	-57.4034528	.	.
	ATE			-55.6255834	.	.
Reading	Unmatched	450.990426	526.960655	-75.9702298	3.46617329	-21.92
	ATT	451.333214	514.575319	-63.2421043	3.65764477	-17.29
	ATU	526.413294	461.216983	-65.1963114	.	.
	ATE			-64.504699	.	.
Science	Unmatched	455.043859	521.854865	-66.8110059	3.61199158	-18.50
	ATT	455.348638	514.308367	-58.9597295	3.72422838	-15.83
	ATU	521.394333	458.503919	-62.890414	.	.
	ATE			-61.4993076	.	.

TABLE 8.17: Matching results - Portugal

Subject	Sample	Treated	Controls	Difference	S.E.	T-stat
Math	Unmatched	452.700683	498.78321	-46.0825264	3.03092922	-15.20
	ATT	452.811471	488.573572	-35.7621004	3.27223281	-10.93
	ATU	498.855342	460.093741	-38.7616012	.	.
	ATE			-37.1076277	.	.
Reading	Unmatched	449.028959	507.883224	-58.8542648	2.89158366	-20.35
	ATT	449.174283	496.48578	-47.3114977	3.09961918	-15.26
	ATU	507.918846	458.5122	-49.4066452	.	.
	ATE			-48.2513468	.	.
Science	Unmatched	456.829256	509.68211	-52.8528536	2.99125833	-17.67
	ATT	456.950104	499.609846	-42.6597423	3.22685823	-13.22
	ATU	509.724083	464.284238	-45.4398456	.	.
	ATE			-43.9068515	.	.

Bibliography

- Use of computers and ICT in schools / PISA

Angrist, J., & Lavy, V. (2002). New evidence on classroom computers and pupil learning. *The Economic Journal*, 112(482), 735-765.

Avvisati, F., Hennessy, S., Kozma, R. B., & Vincent-Lancrin, S. (2013). Review of the Italian strategy for digital schools. *OECD Education Working Papers*, (90), 0_1.

Cheung, A. C., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational research review*, 9, 88-113.

Condie, R., & Munro, B. (2007). The impact of ICT in schools: Landscape review. Becta Research. *Quality in Education Centre*, University of Strathclyde.

De Witte, K., & Rogge, N. (2014). Does ICT matter for effectiveness and efficiency in mathematics education?. *Computers & Education*, 75, 173-184.

Korte, W. B., & Hüsing, T. (2006). Benchmarking access and use of ICT in European schools 2006: Results from Head Teacher and A Classroom Teacher Surveys in 27 European countries. *Empirica*, 1, 0.

Livingstone, S. (2012) Critical reflections on the benefits of ICT in education, *Oxford Review of Education*, 38:1, 9-24.

OECD (2001), "Understanding the Digital Divide", *OECD Digital Economy Papers*, No. 49, OECD Publishing, Paris.

OECD (2015), Students, Computers and Learning: Making the Connection, PISA, *OECD Publishing*.

OECD (2016), PISA 2015 Results in Focus, *OECD Publishing*.

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

OECD (2017), PISA 2015 Technical Report, *OECD Publishing*. (*Final draft)*.

Park, H. (2013). Re-evaluating education in Japan and Korea. *Routledge*.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142-151.

Suppes, P. (1971). Computer-Assisted Instruction at Stanford. *Psychology and Education Series*. Institute for Mathematical Studies in the Social Sciences. Stanford University.

Woessmann, L., & Fuchs, T. (2004). Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school. *CESIFO working paper no. 1321*

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2-3), 114-128.

Zhao, Y., & Lei, J. (2009). New technology. *AERA handbook on educational policy research*, 671-693.

- Propensity score matching

Arceneaux, K., Gerber, A. S., & Green, D. P. (2006). Comparing experimental and matching methods using a large-scale voter mobilization experiment. *Political Analysis*, 37-62.

Arceneaux, K., Gerber, A. S., & Green, D. P. (2010). A cautionary note on the use of matching to estimate causal effects: An empirical example comparing matching estimates to an experimental benchmark. *Sociological methods & research*, 39(2), 256-282.

Blundell, R., Dearden, L. and Sianesi, B. (2005) Evaluating the impact of education on earnings in the UK: models, methods and results from the NCDS. *Journal of the Royal Statistical Society*, Series A 168(3): 473–512.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1), 31-72.

Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association*, 94(448), 1053-1062.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, 84(1), 151-161.

Grilli, L. & Rampichini, C. (2011). Training Sessions on Causal Inference Bristol

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4), 605-654.

Imbens, G. W. (2015). Matching methods in practice: Three examples. Journal of Human Resources, 50(2), 373-419.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Lechner, M. (2008). Matching estimation of dynamic treatment models: Some practical issues. In *Modelling and Evaluating Treatment Effects in Econometrics* (pp. 289-333). Emerald Group Publishing Limited.

Mitra, R., & Reiter, J. P. (2016). A comparison of two methods of estimating propensity scores after multiple imputation. *Statistical methods in medical research*, 25(1), 188-204.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 41-55.

Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: relating theory to practice. *Biometrics*, 249-264.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3), 169-188.

Silverman, B. W. (1986). Density estimation for statistics and data analysis (Vol. 26). *CRC press*.