Is information and communication technology satisfying educational needs at school?

Ferraro, Simona

Tallinn University of Technology

12 April 2018

Is Information and Communication Technology

Satisfying Educational Needs at School?

Simona Ferraro[b]

Abstract

This paper assesses how the integration of ICT in education has affected the mathematics test scores for Italian students measured by the Programme for International Student Assessment 2012 data. The problem of endogeneity that affects survey data in this area, is addressed by applying the Bayesian Additive Regression Trees (BART) methodology as in Cabras & Tena Tena Horrillo (2016). The BART methodology needs a prior and likelihood functions using the Markov Chain Monte Carlo (MCMC) algorithm to obtain the posterior distribution. Controlling for socio-economic, demographic and school factors, the predicted posterior distribution implies an increase, on average, of 16 points in the test scores. The result indicates that the use of ICT at school has a positive and strong impact on mathematic test scores.

JEL Classification: I20, O33, C25

Keywords: ICT, Bayesian Additive Regression Tree, posterior distribution, PISA

[a] The author would like to thank Karin Jõeveer, Kaire Põder, Karsten Staehr, Ingrida Tatolyte and two anonymous referees for very useful comments to earlier versions of the paper. The author would also like to thank INVALSI for providing the dataset.

[b] *Corresponding author:* Simona Ferraro, Tallinn University of Technology, Akadeemia tee 3, 12618 Tallinn, Estonia. Email: simona.ferraro@ttu.ee

1. Introduction

Since the beginning of the 1990s, Information and Communication Technology (ICT) and its impact on students' achievements have interested educators and policy makers. Alike with the aim to improve digital skills among students, the Organization for Economic-Cooperation and Development (OECD 2010, p. 102) has advanced its use with the argument that there is "a significant influence or effect of ICT on the measured or perceived quality of (parts of) education".

The United Nations Educational, Scientific and Cultural Organization (UNESCO) states that "ICT adds value to the processes of learning, and in the organization and management of learning institutions. The Internet is a driving force for much development and innovation in both developed and developing countries" (UNESCO, 2002, p. 9). ICT may indeed be seen as important for the quality of the education systems. Through education, a country creates human capital needed to lead to a higher economic growth (Barro, 2001; Hanushek & Kimko, 2000).

Following the OECD guidelines, European countries has made substantial investments in ICT for educational purposes (OECD, 2015). The European Commission highlighted the use of ICT for work, leisure and communication as among the key abilities and strengths that students need to improve (European Commission, 2006). According to the PISA results from 2009, one computer was available for every two students within schools for most of the OECD countries except for Italy. The disparity within the Italian country is high and only one computer was available up to eight students. In 2012, the gap had decreased and the students-computer ratio was 4.1 to 1 meaning one computer available at school for every four students (Eurydice, 2011 Figure E3, OECD, 2015).

The European Union (EU) has also advised its Member States to invest in digital technologies within their education systems. The Member States agreed to promote the use of new ICT tools within the first cycle of the Strategic Framework for Education and Training known as 'ET 2020' (Eurydice, 2011). This initiative followed the eLearning initiative promoted in 2000 by the European Commission (2000) with the goal to improve the effectiveness of European education systems, and also the competitiveness of the European economy. The integration of ICT at school has seen as a powerful tool to improve technology-related competencies for all students.

The literature regarding the impact of ICT on students' achievements is quite extensive. Several meta-analyses, experimental and parametric studies have been produced but the literature is not unanimous with regard to the effect of ICT on educational outcomes. Cheung & Slavin (2013) provide a meta-analysis study showing that the use of technological applications in education have, in general, a positive impact on students' outcome. Another study from Germany adopting the Programme for International Students Assessment (PISA) survey shows that there is no effect of the use of ICT on PISA test scores (Wittwer & Senkbeil, 2008). The divergent outcomes of those studies suggest that new evidences and approaches are warranted.

This paper assessed whether the use of ICT have had an impact on test scores using large-scale data such as OECD-PISA survey by a new flexible nonparametric model. PISA 2012 survey is a very rich data which also contains questions on the use of ICT among students. As Rojano (1996) states in her work, technology allows students to have the perception of owning the subject. Using a computer with the appropriate software, students can present and observe solutions in real time, for example, how the shape of a geometric object can change.

The paper contributes to the literature by providing new evidences with the application of a new non-parametric methodology compared to what previous researchers have applied (Angrist & Lavy 2002; Cheung & Slavin 2013; Machin et al. 2007). Moreover, students from 32 OECD countries included Italy, who participated in the PISA 2012 Paper-Based Assessment (PBA) were also invited to take a reading and mathematics test on computers[1]. The case of Italy is interesting given the results of Italian PISA 2012 test scores above the OECD average using the Computer-Based Assessment (CBA) compared to the results that students obtained in PISA 2012 using the Paper-Based Assessment (PBA). It is worth to mention that Italian PISA scores using PBA test have usually been below the OECD average in all different set of tests such as reading, mathematics and science (OECD, 2015).

[1] Computer-Based Assessment for PISA 2012 did not include science literacy.

The paper uses a new nonparametric methodology known as Bayesian Additive Regression Trees (BART). The BART model was developed by Chipman et al. (2010) and Hill (2011) and, first applied by Cabras & Tena Horrillo (2016) using Spanish PISA 2012 data. It is a new flexible econometric model which makes it possible to deal with the problem of endogeneity that arises using surveys such as PISA. From an econometric perspective, student and school characteristics may be correlated and the omission of some variables may generate endogeneity bias.

The motivation to adopt the BART model in the field of economics of education is, mainly related to the fact that learning processes are complex, unknown and very heterogeneous. The model relaxes the parametric assumptions and addresses the sample selection problems in survey data as PISA data. The BART model has the aim of providing new results on a single country perspective and stimulate further debates among researchers. The focus, then, on student performance in mathematics is highly correlated to results in reading so that the results presented for mathematic test score can be generalized also to reading.

The rest of the paper is organised as follows. Section 2 reviews the literature, Section 3 presents a review of the Bart Additive Regression Trees and data used in the empirical study, Section 4 provides the main estimation results and robustness check, and Section 5 summarises the results.

2. Literature

The study of the impact of ICT use on mathematics test score in primary and secondary schools has gained interest in the academic literature since the beginning of the '90s. Meta-analyses and single studies have been published but, results are still mixed as several studies have shown (Hatlevik et al. 2015a, 2015b; Rutten et al. 2012; Luu & Freeman, 2011; Tamin et al. 2011; Balanskat et al. 2006; Pedrò, 2006; Hakkarainen et al., 2000; Kulik & Kulik, 1991).

The existing literature can be grouped into three main areas according to the methodology applied in the study. Some studies have adopted parametric models (Machin et al., 2007; Goolsbee & Guryan, 2006, Angrist & Lavy, 2002), a study was conducted through experiment as in Banerjee et al. (2007) while others adopted non-parametric models (Cabras & Tena Horrillo, 2016; Spiezia, 2010; Fuchs & Woessman, 2004). This literature review, however, provides a short survey of the numerous studies on the topic.

Angrist & Lavy (2002) adopt Ordinary Least Squared and the Instrumental Variables (IV) strategy using test scores for Israeli for 1998. They find a negative impact for mathematic test scores after the introduction of computers at school. In the UK, Machin et al. (2007) also using IV strategy show that higher investment in ICT leads to better educational outcomes for reading and science but not for mathematics. Regarding studies using experiment, Banerjee et al. (2007) conduct a randomized experiment in India to study the causal impact of computers on students' performance. They compare the change in the test scores among students who received the treatment and students who did not receive it represented by the use of a computer. They find that students who were able to use a computer have also higher mathematic test score compared to their peers.

Among some studies that have applied non-parametric methods and PISA data, there are works by Spiezia (2010), Fuchs & Woesmann (2004), Shewbridge et al. (2005). Using PISA 2006 for science score in OECD 33 countries, Spieza (2010) shows a positive correlation between the availability of computers at school and school performance. He estimates an endogenous treatment model where the frequency of computer use is modelled on specific students' characteristics. In their study, Fuchs & Woessmann (2004) control for numerous variables with a known impact on achievement using PISA 2000 and two stage least squares. They show that there is a positive correlation of the home computer use but this effect is almost neutral or even negative for the use of computer at school.

In their studies, Shewbridge et al. (2005) find no effect of the use of ICT at school and PISA 2003 test scores such as reading, mathematics and science. A new flexible non-parametric approach from the Bayesian family, has been employed with PISA data by Cabras & Tena Horrillo (2016). They study the causal impact of ICT on educational outcomes adopting a new model from the Bayesian models for Spanish PISA 2012 data that represent a general survey for the whole student population including a rich collection of information on individual, family and school levels.

The BART model allows them to overcome the endogeneity problem that arises from the survey and using some control variables, their findings show that the use of ICT has a strong positive effect on students' achievements. This paper follows the methodology in Cabras & Tena Horrillo (2016) for Italian data.

In Italy, the investment in ICTs and its introduction at all levels of compulsory education, is crucial for the development of digital skills (Annali della Pubblica Istruzione, 2012). The Ministry of Education have released a survey known as 'Teaching Multimedia Equipment Survey' to individuate technological instruments adopted by schools such as the use of the Internet, amount and speed of Internet connections, ratio of classrooms equipped with wireless connectivity, total number of computers (desktop and laptop), mobile devices. The data are available and uploaded on the "Scuola in Chiaro platform". Despite of the geographical gap between North and South, the Italian National Statistics Office (Istat, 2015) showed that the gap is also present in digital infrastructures but, it underlines that during last years the Southern regions received specific funds from the National Operational Program showing a reduction in the gap. A detailed study of the Italian Strategy for Digital Schools can be found in Avvisati et al. (2013).

The Italian literature in education has been focused mainly on the Family Background effects, school level peer effects and also on the causes of the regional disparities using PISA data or INVALSI administrative data (Checchi 2004; Bratti et al. 2007; Montanaro 2007, Agasisti & Vittadini 2012). Focusing on both the use of computer at home and at school with PISA 2006 data, Ponzo (2011) shows that students' achievements are negatively affected when students use computer at school compared to using a computer at home. A more recent study that focuses on the impact of ICT on students' outcome which includes Italian data is by Agasisti et al. (2017). They employ data from the OECD-PISA 2012 for 15 European countries. Despite they focus mainly on the effect of using ICT at home for school related tasks, they also show that for higher values of ICT used at school, there is a decrease in the test score.

3. Research Method

This section discusses data collection and the research design represented by the econometric model for estimating the causal effect of some treatment variables identified with the use of computer or laptop or tablet at school, on students' mathematic test scores. It also includes the data analysis and all variables used in the study.

3.1 Participants

This article uses the fifth wave of PISA survey conducted by the OECD in 2012. PISA survey is administrated by the Italian National Evaluation Committee (INVALSI) and subsequentially is elaborated by the OECD. The PISA survey is a cross-national survey, carried out every three years and since 2000, its main goal is to assess 15 year-old students' performance in reading, mathematics and science literacy as well as problem-solving skills.

Since 2009, the survey contains a questionnaire on students' familiarity with ICT where students give information on which kinds of technology they have at their disposal at home and also at school; whether they use them and how often they use them and for what purposes. The survey also contains questions for self-assessment, in other words it askes the level of proficiency and confidence of students using a ICT tool. The database contains detailed information on students' characteristics as well as on family and schools characteristics.

The Italian PISA results are interesting to study because, in 2012, the survey was also conducted by CBA test and Italian students improved their test scores compare to the PBA test). Results from PISA 2012 survey showed an improvement of students' scores: 504 points for reading and 499 for mathematics in the CBA test against 490 points for reading and 485 points for mathematics in the PBA test (OECD, 2015). It is, hence, interesting to conduct this study on Italian case. A description of PBA and CBA scores are presented in Table 1.

Table 1: Italian PISA 2012 test scores

	Math	SE	Reading	SE	Science	SE
Paper-Based Assessment	485	2.0	490	2.0	494	1.9
OECD Average	494	0.5	496	0.5	501	0.5
Computer-Based Assessment	499	4.2	504	4.3	---	---
OECD Average	497	0.7	497	0.7	---	---

Source: OECD-PISA 2012
Notes: In 2012, the Computer-Based Assessment (CBA) was only for reading, mathematics and problem solving.

3.2 Research design (Review of Bayesian Additive Regression Trees)

Causality defines the causal relationship in terms of potential outcome frameworks for describing what would happen to a given individual in a hypothetical comparison of alternative scenarios. In the first case, there is the factual situation while in the other case the counterfactual situation. In this paper, counterfactual situations are operationalized by using the notation suggested by Rubin (1978). The potential outcomes are:

$$\text{Potential outcome} = \begin{Bmatrix} Y_{1i} = outcome\ for\ i_{th}\ student\ if\ treated \\ Y_{0i} = outcome\ for\ the\ i_{th}\ student\ if\ not\ treated \end{Bmatrix} \quad (1)$$

The dependent variable is the OECD-PISA mathematic test score, while the treatment is a binary variable expressed by $Z = 1$ whether a student uses a computer, a laptop or a tablet at school or $Z = 0$ otherwise. The potential outcome Y_{1i} measures the mathematics test scores where the subscript 1 indicates whether the computer, laptop, tablet exist and are in use at school for the individual *i* while Y_{0i} where the subscript 0 indicates whether the computer, laptop, tablet exist and are not used at school. In other words, Y_{1i} and Y_{0i} are the potential outcomes for individual *i* and the casual effect of the treatment variable Z for using ICT at school, on test scores.

The observed outcome Y_i can also be expressed by $Y_i = Y_{0i} + (Y_{1i} - Y_{0i})\ Z_i$ in terms of potential outcomes and treatment effect, where Z_i is a treatment dummy variable. In studies as OECD-PISA survey, scholars face the endogeneity problem because potential results are not independent from the treatment variable. An endogeneity problem may arise, for example, when families with high socio-economic status enrol their children in schools that have better and well-equipped informatics rooms compared to families with low socio-economic status who can decide to enrol their children in schools that invest less in informatics infrastructures. Therefore, it is more likely that family decisions and socio-economic status affect students' test scores.

To overcome the endogeneity problem and assuming independency among outcomes and treatment variables, several covariates should be included in the model controlling for some individual characteristics. The recent use of non-parametric methodologies like the BART model discussed in sub-section 3.2, avoid to have several different models to capture the endogeneity as classical approaches such as linear regression models or propensity score do not allow, indeed, to overcome the problem as treatment and not treatment are not observable for a specific characteristic of the individual indicated with X. Given that treatment and no treatment are not observable for the same value of X, the estimation of the score assigned to each individual becomes difficult and as alternative, the nonparametric methods are more flexible compared with linear models.

The Average Treatment Effect (ATE) is computed as the difference between $Y_{1i} - Y_{0i}$ cannot be computed as direct measure because Y_{1i} and Y_{0i} are not directly observable. For a given treatment and control condition, each student i can have two potential outcomes: $Y(0)$ and $Y(1)$ where $Y(Z = 1) = Y(1)$ if students receive the treatment while $Y(Z = 0) = Y(0)$ otherwise. The ATE equals $E\ (Y(1) - Y(0)\ )$ and it defines the expected value with respect to the probability distribution of the dependent variable for all the individuals. The variables of interest is the expected value of potential outcomes conditional to the treatment $E(Y(1) - Y(0)|Z = 1)$. To address a possible bias, the model uses the Conditional Independence Assumption (CIA) conditional on observed individual characteristics indicated by X_i.

Looking at individuals with the same characteristics, $\{Y_{1i}, Y_{0i}\}$ and the treatment Z_i, the dependent variable Y is conditional independent:

$$\{Y_{1i}; Y_{0i}\} \text{ independent of } Z_i, \text{ conditional on } X_i \qquad (2)$$

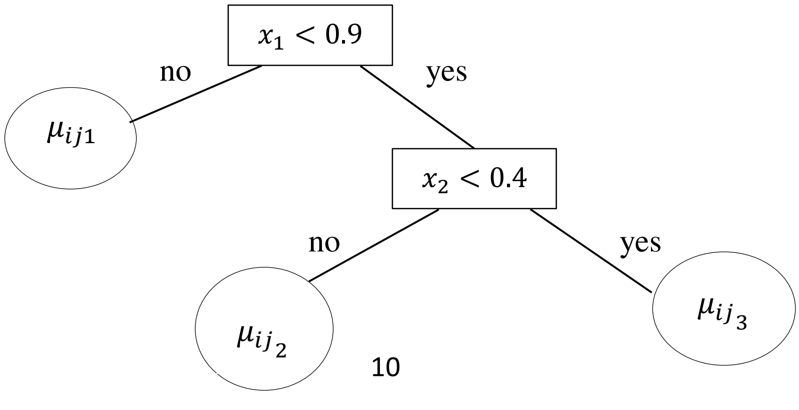
The Bayesian Additive Regression Trees (BART) provides a framework for flexible nonparametric modelling of the relationships of covariates to outcomes and it is a tree-based variable selection making use of the internals of the decision tree structure.

3.3 Estimation of the model BART

Decision tree ensembles have become a popular tool for obtaining high quality predictions in nonparametric regression problems, also motivated by the success of methodological approaches such as boosting (Chipman et al. 2010; Denison et al. 1998). They use an algorithm to learn the relationship between the response and its predictors (Breiman, 2001) assuming that the data-generating process is complex and unknown. In this framework, the approach of the BART model allows to estimate the response outcome and the counterfactual result using an extension of a non-parametric Bayesian model that performs conditional inference without making any pre-assumption on the distribution as classical inference does.

The BART model consists of a collection of regression tree models. Considering y_i as the outcome and x_i as a vector of covariates where their relationship is given by the function $y_i = g(x_i; T, M) + \varepsilon_i$ where $g(x_i; T, M)$ is a binary tree function, T indicates the tree structure that consists of two sets of nodes: an interior and a terminal node and, a branch decision rule at each interior node. The branch decision rule is typically a binary split based on a single component of the covariate vector. The second tree component is $M = \{\mu_1, \dots, \mu_b\}$ is made up of the function values at the terminal nodes. An example is provided in Figure 1.

Figure 1: Example of single binary tree with branch decision rules (circles) and terminal nodes (rectangular)



This paper follows the application of the BART methodology for Spanish PISA data in the study by Cabras & Tena Horrillo (2016). Their study is the first that applies a non-parametric model within the framework of Bayesian models in educational studies. The aim of BART is to estimate the posterior probability distribution of the causal effect conditional to some covariates π (ATE|X) using its flexibility in high non-linear response surfaces even with a large number of predictors (with great out-of-sample prediction properties).

For this BART model, there are a sum of trees with a prior distribution over the depth of the splits and the values at the leaf nodes. The a sum of trees is fitted in the context of the rest and on the iterative algorithm and each tree is modified one by one based on the residuals from the generation of previous trees (unlike random forests, where each tree is independent). This means that there is an informative prior and allows BART to better captures additive effects.

Formally, all observations begin in a single root node and then, the root node's splitting rule is chosen by the algorithm and consists of a splitting variable x_k and a split point c. The observations in the root node are split into two groups, based on whether the splitting variable is greater or smaller than the split point $x_k \geq c$ or $x_k < c$. The two groups become a right daughter node and a left daughter node while within each of the two nodes, additional binary splits can be chosen.

The equation for the basic BART model also defined as likelihood function is the following:

$$Y = \sum_{j=1}^{m} g\left(x_k, z; T_j, M_j\right) \tag{3}$$

where $g(x_k)$ is a Bayesian decision tree model as described in Chipman et al. (2010, 1998) with x_k as splitting variables, z is the treatment effect that belongs to the individual whose response is Y and have the error term normally distributed $\varepsilon \sim N(0, \sigma^2)$ where σ^2 the residual variance. The term T_j refers to decision tree where *j* refers to the number of trees which goes from 1 to *m,* where *m* is the total number of trees in the model while M_j is the function values at the terminal nodes.

The Additive Regression Trees employs an ensemble of such trees in an additive fashion, that is, it is the sum of m trees where m is typically large such as 200, 500, or 1000. The model is fitted via a back-fitting Gibbs sampler that draws from the joint posterior distribution of all the trees and terminal node parameters and the standard deviation, given the data (Chipman et al., 2010). Each tree T_j is iteratively fitted and based on the residuals generated from the previous trees, at the current iteration of the Gibbs sampler until a predetermined number of iterations is reached. The prior on T_j and M_j strongly favours small trees and leaf parameters that are near zero, constraining each term in the sum to be a "weak learner".

Starting from the root node, the probability that a node at depth d splits (is not terminal) is given by $\alpha(1 + d)^{-\beta}$ where $\alpha \in (0.1), \beta \in [0, \infty)$ where d is the depth of internal node i and, α, β are parameters that determine both the size and shape of the trees. This paper employs the standard values with $\alpha = 0.95$ and $\beta = 5$ as indicated in Chipman et al. (2010) and in Cabras & Tena Horrillo (2016). Such values assure that trees do not grow too much and each tree with more than 5 terminal nodes has a probability of 3 per cent. The model also uses the Markov Chain Monte Carlo (MCMC) and the Metropolis Hastings within Gibbs for simulating samples from the posterior distribution with a non-excessive computational effort. For this study, $m = 200$ trees and 5000 MCMC steps after an initial burn-in of 1000 steps are used. Interactions are estimated from the data by the 500th tree and they are not specified in the model a priori. For the estimation, R software was used and the package "bartMachine" recently developed by Kapalner & Bleich (2016).

3.4 Data analysis

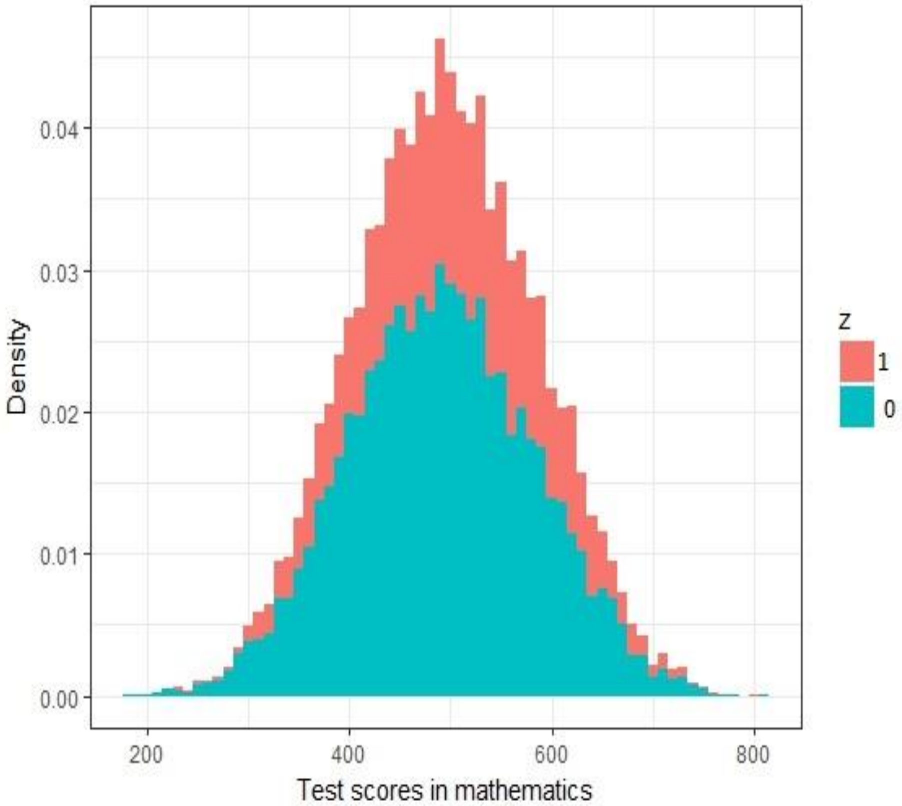
The sample consists of 21,520 observations and 925 variables. The final PISA sample was chosen randomly and the selection probabilities of students vary so weights must be used to be sure that the sample represents correctly the full PISA population (OECD, 2014). The final sample consists of 21,520 observations and 17 variables chosen for the study. It is a sub-sample created to estimate the prediction model for the dependent variable that it is the PISA test score in mathematics defined as plausible values (pv1math).

The estimation uses variables related to students' characteristics such as gender, relative age related to whether the student is born before the first half of the year (before June) or after, whether the student attended or not the pre-primary school, the immigration status, family structure and how much time a student spend using internet (time internet). The socio-economic status of students is also included and it is expressed in the ESCS index constituted of several indicators: International Socio-Economic Index of Occupational Status (ISEI), the Highest level of education of the student's parents (HISCED), converted into years of schooling. It also includes the index of family wealth (WEALTH), the index of Home Educational Resources (HEDRES) and, the index of possessions related to classical culture in the family home (Home Possession).

Other variables are included to control school characteristics such as the quality of educational resources at school expressed by the variable school_resources, the student-teacher ratio, whether the school is a public or private school. Then, the index of availability of computers with the variable computer_ratio obtained by dividing the number of computers at school by the number of students at school, class size related to the number of students in each class and mathematics teacher-students ratio. Descriptive statistics are summarised in Table 2 in Appendix A.

The treatment variable is the use of computer, laptop, tablet at school indicated by the variable "Treatment" that has value 1 if the student use, at least, a computer, a laptop or a tablet at school. The treatment variable is used to compute the causal effect on the dependent variable, the mathematic test score. The sample histogram for the dependent variable (math test score) for students who use or not use a computer, laptop, tablet (treatment variable) is shown in Figure 2.

Fig. 2: Histogram for distribution of PISA scores conditional to the treatment *Z*



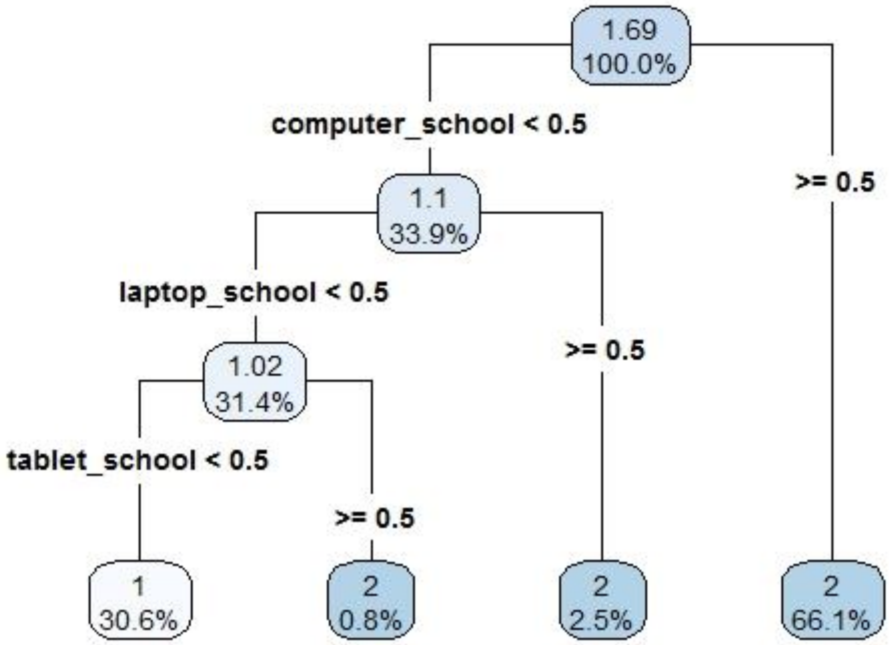
Source: OECD-PISA data 2012 for Italy
Notes: Author's calculation. The treatment Z = 1 indicates students who use a computer, laptop or tablet at school, while Z = 0 indicates students who do not use a computer, laptop or tablet.

The histogram shows the distribution of the use of computer, laptop or tablet among students at school. The treatment variable $Z = 1$ indicates 14,937 students who use it and 6,583 students who do not use it. Therefore, 14,937 out of 21,520 students use ICT and it is an unbalanced sample that classical parametric approaches can estimate without problems.

4. Estimation results

This section presents the main results of the impact of the ICT on Italian mathematic test scores using PISA 2012 data and the flexible BART model discussed in Section 3. After the construction of trees, the fitted values are assigned to each terminal node. The fitted values will be the average of response values for the regression tree and the majority class for the node in the classification trees. Figure 3 shows three steps in the growth of a classification tree for response y with levels "0" and "1" and predictor *Z*.

Figure 3: Three Steps in the growing process of a classification tree

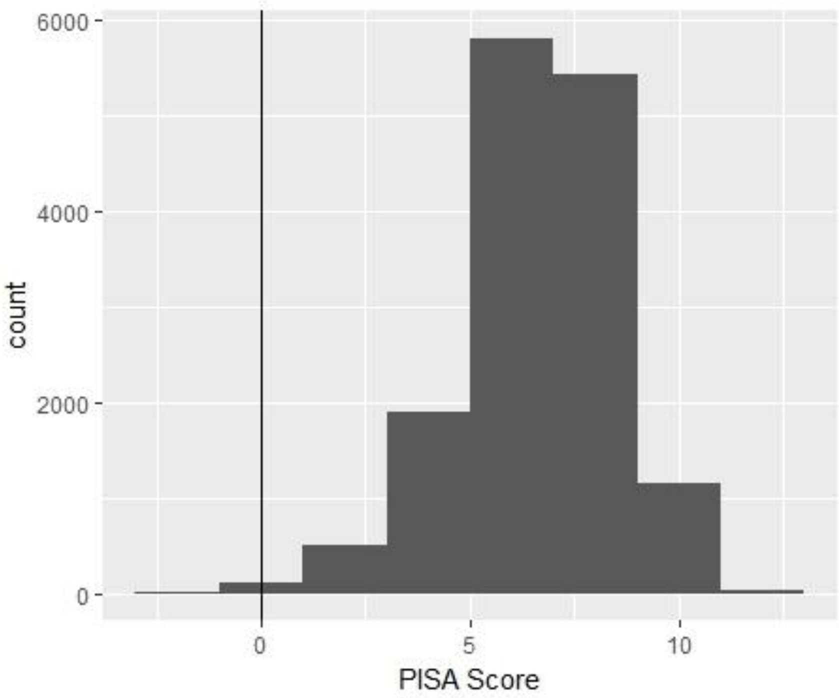


Notes: Author's own calculation.
Source: PISA 2012 data

The classification tree above shows how the root - that is the starting point of the tree - is splitted. If half of students use a computer at school (>=0.5) the predicted number of test scores will be 2 with probability of 66.1%. For the other half of students who do not use a computer at school, there is no final prediction but rather there is another split. The process continues until there are no splitting points. The end part of the tree consists of "leaves". In between root and leaves, there are decision nodes from where new splits are generated. The percentages represents the percentage of the total sample which must be 100%.

The posterior distribution of the marginal causal effect in Figure 4 is derived by the simulated differences between the mean of the posterior predictive distribution for students that use or not use a computer or laptop, or tablet at school at school. The approximation is generated by means of MCMC draws, of the posterior distribution $\pi(ATE|D)$. The posterior distribution for the Italian case shows a positive effect of 99%. The size of the posterior distribution is given by the ratio of the posterior probability and its magnitude which indicates that it is on average, 16 times more likely that ICT, such as the use of a computer or a laptop or a tablet at school, has a positive impact on educational outcomes.

Figure 4: Approximation of the posterior distribution for Italy



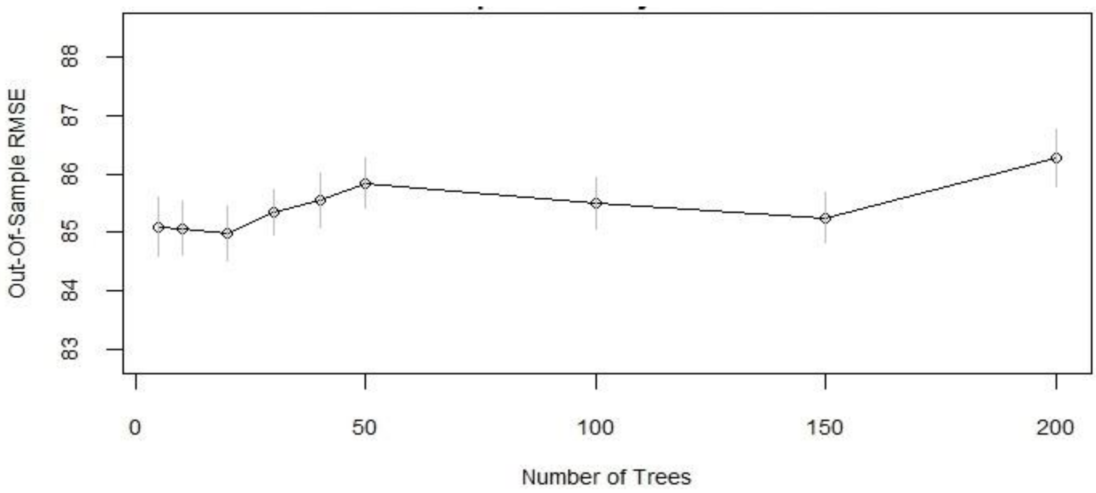
Source: OECD-PISA data 2012
Notes: Elaboration by author using R software

In particular, taking a closer look at the confident interval at 99%, results also show that the effect of ICT is within the interval 2.97 and 9.23 meaning that the effect is substantially strong positive. For the immigration status, meaning whether ICT is more beneficial for native or immigrants students, estimating the ATE and the posterior distribution for both of them the effect is around 99.8% for non-native students and 99.7% for native students. Students who are not natives have also beneficial effects from the use of ICT. It is possible to say that the ICT has anyway a strong impact on immigrants and may be helpful for them for filling some gaps.

The final model aggregates four post-burn-in chains for the four cores indicated in the parameter that yields the 1,000 total post-burn-in-samples. This gives the drawback of effectively running the burn-in serially and add the benefit to reduce auto-correlation of the sum-of-trees samples in the posterior distribution since the chains are independent giving a greater predictive performance. The pseudo-R^2 for in-sample is

0.53 and higher compared to 0.24 of the parametric study adopting the Ordinary Least Squared (Table A2 – Appendix A). Figure 5 illustrates that predictive performance levels off around m = 20 with an improvement with the further trees.

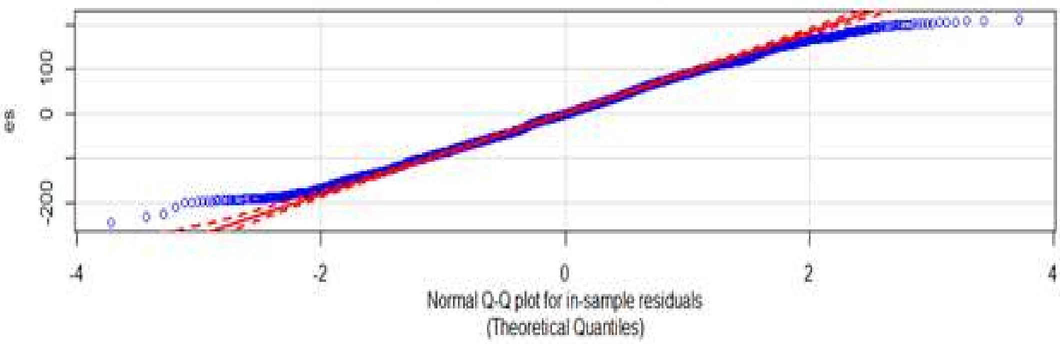
Figure 5: Out-of-Sample predictive performance by number of tree



Source: OECD-PISA 2012
Notes: The starting point is set to *m=200* as in Chipman et al. (2010) and not *m*=500 as in Cabras and Tena Horrillos (2016) to reduce computational time and memory requirements. Performance results are very similar.

The summary of the posterior distribution in Figure 6 shows the p-value for Shapiro-Wilk test of normality of residuals. Figure 5 displays that the predictive performance levels are off around 50 trees and there is a stationary trend.

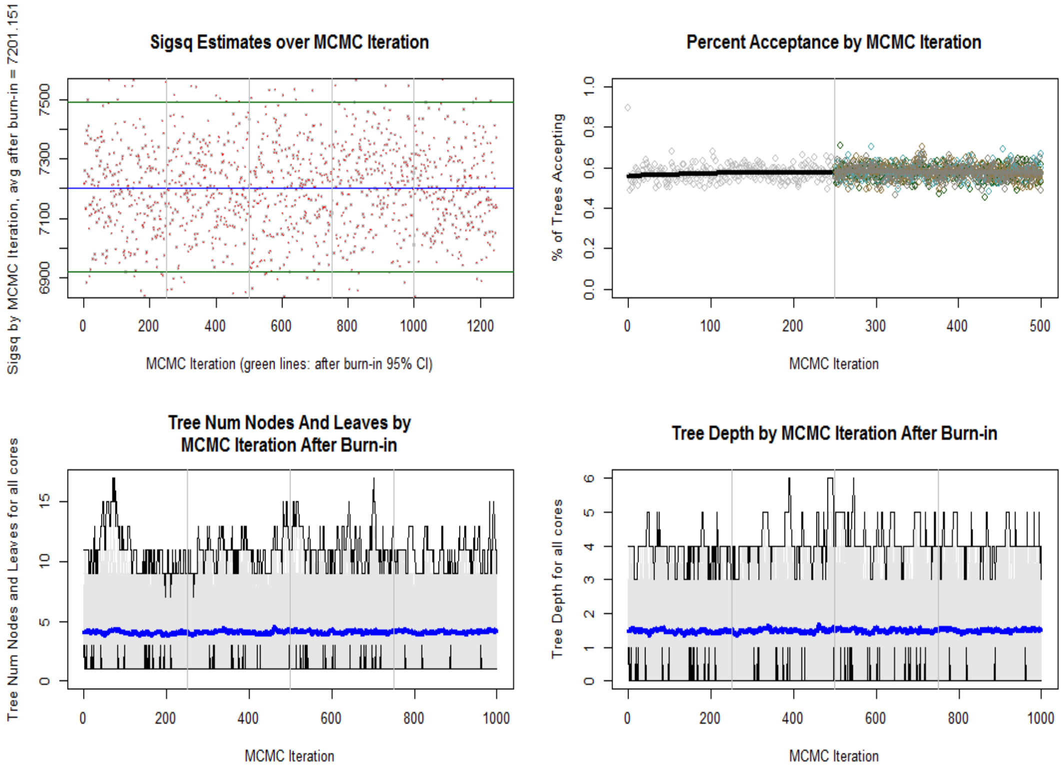
Figure 6: Assessment of Normality



Notes: Author’s calculation

Figure 6 shows that the assumption of normality is not violated. To check also the convergence of the Gibbs sampler, Figure 7 displays four plots which features the convergence diagnostics.

Figure 7: Convergence diagnostic

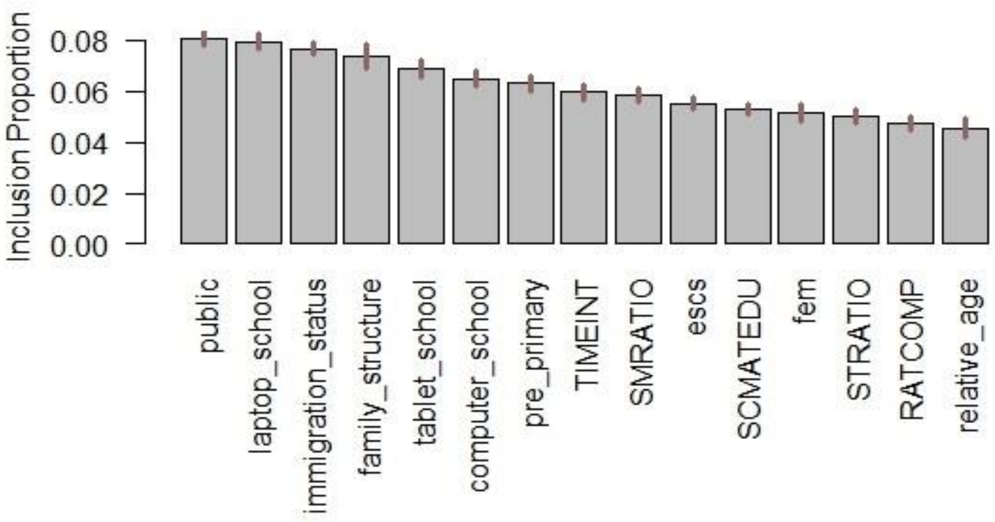


Notes: Author's calculation

The first panel on the top-left is the sigma-squared by MCMC iteration and the plot shows five boxes. The first box on the left indicates burn-in from the first computing core's MCMC chain while the following four plots show the post-burn-in iterations from each of the four computing cores. The second plot on the top-right indicates the percent acceptance of Metropolis-Hastings proposals for all trees where each point represents one iteration. It is possible to see two boxes: the box on the left illustrates burn-in iterations and points after illustrate post-burn-in iterations. For last two plots on the bottom, the plot on the bottom-left shows the average number of leaves across the m trees by iteration while the plot on the bottom-right shows the average tree depth across the m trees by iteration. It is visible that the model has burned-in quite nicely and each plots exhibits a stationary process.

In conclusion, it is possible to check which variables are the most important in the model counting how many times a variable appear in a tree indicating which variables have a more important role in affecting students' results. Figure 8 indicates the average variable inclusion proportions.

Figure 8: Average inclusion proportion



Notes: Author's calculation using R software.

This figure shows the results after assessing the splitting rules in the *m* trees across the post-burn-in MCMC iterations. This process is also known as inclusion proportions (Chipman et al. 2010) and it represents for a given predictor or covariate, the proportion of times that the variable has been chosen as a splitting rule out of all splitting rules among the posterior draws of the sum-of-trees model.

For this study, the variable which appears several times and is the most important for explaining the response is the variable determining whether school is a public or private school. In this study, the variable assumes dummy characteristics with public school taking value 1. It is possible to conclude that public school variable has more weight in predicting results in test score. A robustness check is presented comparing the BART model with the traditional parametric model such as the linear regression. Estimation results of the linear regression are presented in Table 2.A in Appendix A. Results indicate that the use of ICT at school increase PISA test scores for 6.5 points but the goodness-of-fit model is low. The interpretation for the low value compared to what the Bayesian analysis computed, shows the low power of the linear model to fit all variables.

The analysis and interpretation of the results from the Bayesian analysis suggest the positive impact of the ICT on students' test scores and the predictive power of the model to explain the causal effect.

5. Conclusion

This paper studies the impact of ICT expressed as the use of a computer, laptop or tablet at school on the mathematic test scores for Italian students. The Italian PISA 2012 data was employed and the BART model was applied. The Italian case was an interesting case to study after the improvement of Italian students' scores in reading and mathematics when the Computer-Based Assessment test was carried out by OECD, in 2012. Italian PISA scores have usually been below the OECD average as confirmed by OECD reports (2007, 2010, 2016). The BART model was applied as flexible Bayesian methodology with some advantages compared with other classical parametric model such as: (i) overcomes the problem of endogeneity and (ii) uses less assumptions in the specification of the model. Moreover, as Cabras and Tena Horrillo (2016) say in their study, the interpretation of the coefficient in the Ordinary Least Squares is challenging because of the difficulty to introduce in the model, all relevant covariates with all the interactions.

The analysis has shown that computer use does increase student performance. This study is also innovative in that because it moves beyond the descriptive analysis of the country. The study applies a different econometric model that is not based on parametric assumptions. As in Cabras & Tena Horrillo (2016) who used BART model for Spanish PISA 2012 data showing a positive effect of the use of ICT on Spanish students' outcomes, the posterior distribution ATE for Italian data in Figure 4 underlined the positive and strong effect of the treatment variable – use of computer or laptop or table at school - on mathematic test scores. The impact can be computed in almost 16 times more likely for students who use ICT to improve their test score. In this respect, results from this study are in line with those in Cabras & Tena Horrillo (2016).

The paper shows that using ICT at school leads to better learning and knowledge acquisition among students and leads to better results among the students' mathematics scores. The non-parametric analysis as the Bayesian analysis is able to overcome the issue in PISA when the number of potential confounding variables is large. As it does not require any subjective decision by the scholar expect for the indication of the treated variable, BART allows to be also implemented also in different contexts.

Analysis of normality but also the analysis of the converge diagnostics showed that BART and the burned-in MCMC iterations provided a good approximate posterior distribution. A robustness check using the parametric model is also presented showing the effect of the treatment, on the sample of Italian students. The positive sign of the treatment variable is statistically significant at 1 per cent level but the coefficient of determination is lower compared to what the Bayesian analysis showed. However, as Cabras and Tena Horrillo (2016) pointed out, the interpretation of the coefficient in the Ordinary Least Squares is challenging because the difficulty in introducing all relevant covariates with all the interactions in the model. The BART model seems, hence, an effective model for causal inference as Hill (2011) showed as there is no need to estimate several models as traditional parametric analysis such as propensity score matching can require. However, BART model can be demanding in terms of computational algorithm and there is a need for further applications and improvement of the model using also different data.

References

Agasisti Tommaso, Maria Gil-Izquierdo and Seong Won Han (2017). "ICT use at home for school-related tasks: What is the effect on a student's achievement? Empirical evidence from OECD PISA data", Munich Personal RePEc Archive (MPRA) Paper, no. 81343.

Agasisti Tommaso & Giorgio Vittadini (2012). "Regional economic disparities as determinants of student's achievement in Italy", *Research in Applied Economics,* vol. 4, no. 2, pp. 1-33.

Angrist, Joshua & Victor Lavy (2002). "New evidence on classroom computers and pupil learning", *Economic Journal*, vol. 112, no. 482, pp. 735–765OECD.

Annali della Pubblica Amministrazione (2012). "Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione" [Annals of Public Administration (2012): Guidelines for the curriculum of pre-primary education and the first cycle of school education], Special Issue. Milano: Le Monnier

Avvisati Francesco, Sara Hennessy, Robert B. Kozma and Stéphan Vincent-Lancrin (2013). "Review of the Italian Strategy for Digital Schools", OECD Education Working Papers, no. 90, OCED Publishing, Paris

Balanskat, Anja, Roger Blamire and Stella Kefala (2006). The ICT impact report. A review of studies of ICT impact on schools in Europe. European Commission.

Banerjee, Abhijit, Shawn Cole, Esther Duflo and Leigh Linden (2007). "Remedying education: Evidence from two randomized experiments in India", *Quarterly Journal of Economics*, vol. 122, no. 3, pp. 1235-1264.

Barro, Robert J. (2001). "Human capital and growth", *American Economic Review*, vol. 91, no. 2, pp. 12–17.

Bratti Massimiliano, Daniele Checchi and Antonio Filippin (2007). "Territorial Differences in Italian Students' Mathematical Competences: Evidence from PISA, *Giornale degli Economisti e Annali di Economia,* vol. 3, no. 66, pp. 299-335.

Breiman Leo (2001). "Random Forests", *Machine Learning,* vol. 45, no. 1, pp. 5-32.

Cabras Stefano & José de Dios Tena Tena Horrillo (2016). "A Bayesian non-parametric modelling to estimate student response to ICT investment", *Journal of Applied Statistics*, vol. 43, no. 12, pp. 1-16.

Checchi Daniele (2004). "Da Dove Vengono le Competenze Scolastiche? L'indagine PISA 2000 in Italia", *Stato e Mercato,* vol. 72, pp. 416-453.

Cheung Alan C. & Robert E. Slavin (2013). "The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis, *Educational Research Review,* vol. 9, pp. 88-113.

Chipman, Hugh A., Edward I. George and Robert E. McCulloch (2010). "BART: Bayesian Additive Regression Trees", *The Annuals of Applied Statistics*, vol. 4, no. 1, pp. 266–298.

Chipman, Hugh A., Edward I. George and Robert E. McCulloch (1998). "Bayesian CART Model Search", *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 935–948.

Denison David G.T., Bani K. Mallick and Adrian F.M. Smith (1998). "A Bayesian CART Algorithm", *Biometrika,* vol. 85, no. 2, pp. 363-377.

Eurydice (2011). Key Data on Learning and Innovation through ICT at School in Europe. 2011 Edition. Brussels, Eurydice.

European Commission (2000). Communication from the Commission-e-Learning – Designing tomorrow's education. COM(2000) 318 final.

Fuchs Thomas & Ludger Woessmann (2004). “Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school”, *Brussels Economic Review,* vol. 47, no. 3-4, pp. 359-385.

Goolsbee Austan & Jonathan Guryan (2006). “The Impact of Internet Subsidies in Public Schools”, *The Review of Economics and Statistics,* vol. 88, no. 2, pp. 336-347.

Hakkarainen, Kai, Liisa Ilomäki, Lasse Lipponen, Hanni Muukkonen, Marjanna Veermans, Taneli Tuominen, Minna Lakkala and Erno Lehtinen (2000). “Students’ skills and practices of using ICT: results of a national assessment in Finland”, *Computers & Education*, vol. 34 no. 2, pp. 103-17.

Hanushek, Erik A. & Dennis D. Kimko (2000). “Schooling, Labor Force Quality, and the Economic Growth”, *American Economic Review,* vol. 90, no. 5, pp. 1184–1208.

Hatlevik Ove E., Greta B. Guðmundsdóttir and Massimiliano Loi (2015a). **“**Digital diversity among upper secondary students: A multilevel analysis of the relationship between cultural capital, self-efficacy, strategic use of information and digital competence”, *Computers & Education*, vol. 81, pp. 345-353.

Hatlevik Ove E., Geir Ottestad and Inger Throndsen (2015b). **“**Predictors of digital competence in 7th grade: A multilevel analysis”, *Journal of Computer Assisted Learning*, vol. 31, no. 3, pp. 220-231.

Hill, Jennifer H. (2011). “Bayesian Nonparametric Modeling for Causal Inference”, *Journal of Computational and Graphical Statistics,* vol. 20, no. 1, pp. 217–240.

Istat (2015). Rapporto annuale 2015. La situazione del Paese [Annual Report 2015. The situation of the country] Roma: Istat.

Kapalner Adam & Justin Bleinch (2016). “bartMachine: Machine Learning with Bayesian Additive Regression Trees”, *Journal of Statistical Software,* vol. 70, no. 4, pp. 1-40.

Kulik Chen-Lin C. & James A. Kulik (1991). “Effectiveness of computer-based instruction: An update analysis” *Computers in Human Behavior*, vol. 7, no. 1-2, pp. 75-94

Luu, King & John G. Freeman (2011). “An analysis of the relationship between information and communication technology (ICT) and scientific literacy in Canada and Australia”, *Computers & Education*, vol. 56, no. 4, pp. 1072-1082.

Machin, Sthephen, Sandra McNally and Olmo Silva (2007). “New technology in schools: Is there a payoff?”, *Economic Journal*, vol. 117, no. 522, pp. 1145–1167.
Montanaro Paolo (2007). “La Qualità dell’Istruzione Italiana” Mimeo.

OECD (2016). "PISA 2015 Results in Focus", *PISA in Focus*, No. 67, OECD Publishing, Paris.

OECD (2015). Students, computers and learning: Making the connection, PISA, OECD Publishing.

OECD (2014). PISA 2012 Technical Report. PISA, OECD Publishing.

OECD (2010). Are new millennium learners making the grade? Technology use and educational performance in PISA. OECD Publishing, Paris

OECD (2010). PISA 2009 Results: Executive Summary

OECD (2007). PISA 2006: Science Competencies for Tomorrow’s World Executive Summary.

Pedro´, Francisco (2006). The New Millennium Learners: Challenging Our Views on ICT and Learning, OECD-CERI, Paris.

Ponzo, Michela (2011). “Does the way in which students use computers affect their school performance?” *Journal of Economic and Social Research*, vol. 13, no. 2, pp. 1–27

Rojano, Teresa (1996): "The role of problems and problem solving in the development of algebra. In N. Bednarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 137–145). Dordrecht, Netherlands: Kluwer Academic Publishers.

Rubin, Donald B. (1978). "Bayesian Inference for causal effects: The role of randomization", *Annals of Statistics,* vol. 6, no. 1, pp. 34-58.

Rutten, Nico, Wouter R. van Joolingen and Jan T. van der Veen (2012). "The learning effects of computer simulations in science education", *Computers & Education*, vol. 58, no. 1, pp. 136-153.

Shewbridge, Claire, Miyako Ikeda and Andreas Schleicher (2005). "Are students ready for a technology rich world? What PISA studies tell us", OECD Paris.

Spieza, Vincenzo (2010). "Does computer use increase educational achievements? Student-level evidence from PISA", *OECD Journal*: *Economic Studies*, 2010, pp. 1–22.

Tamim, Rana M., Robert E. Bernard, Eugene Borokhovski, Philip C. Abrami and Richard F. Schmid (2011). "What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study", *Review of Educational Research*, vol. 81, no. 1, pp. 4-28.

UNESCO (2002). "Information and Communication Technology in Education. A Curriculum for Schools and Programme of Teacher Development", UNESCO, Paris.

Wittwer Jorg & Martin Senkbeil (2008). "Is students' computer use at game related to their mathematical performance at school?, *Computers & Education,* vol. 50, no. 4, pp. 1558-1571.

Appendix A

Table A.1 Descriptive Statistics

	Use of computer = 1					Use of computer = 0				
	Obs	**Mean**	**S.D.**	**Min**	**Max**	**Obs**	**Mean**	**S.D.**	**Min**	**Max**
Relative age	14,937	0.50	0.50	0	1	6,583	0.49	0.50	0	1
Gender	14,937	0.49	0.50	0	1	6,583	0.53	0.50	0	1
Attended pre-primary school	14,937	0.96	0.19	0	1	6,583	0.97	0.17	0	1
ESCS	14,937	-0.08	0.93	-0.75	2.70	6,583	0.12	0.97	-4.7	2.7
Family structure	14,937	0.90	0.30	0	1	6,583	0.90	0.30	0	1
Immigration status	14,937	1.10	0.42	1	3	6,583	1.09	0.40	1	3
Time spent on internet (TIMEINT)	14,937	40.44	33.68	0	206	6,583	37.46	32.94	0	206
Resources at school (SCMATEDU)	14,937	0.09	0.88	-3.59	1.98	6,583	-0.12	0.87	-3.59	1.98
Teacher-student ratio (STRATIO)	14,937	9.62	3.82	0.63	72.54	6,583	10.78	3.30	0.63	28.18
School type	14,937	0.97	0.16	0	1	6,583	0.98	0.13	0	1
Computer-student ratio (RATCOMP)	14,937	0.59	0.45	0	4.5	6,583	0.41	0.33	0	4.5
Mathematic teacher-student ratio (SMRATIO)	14,937	92,74	38.82	5.65	651	6,583	90.63	34.90	5.65	335

Source: OECD-PISA 2012

Table A.2. Linear Regression analysis

Variables	(2.1)
Relative age	6.891*** (1.061)
Gender (female=1)	-22.290*** (1.069)
Pre-primary school (yes=1)	33.994*** (2.933)
ESCS	19.831*** (0.579)
Family structure (nuclear=1)	0.172 (1.775)
Immigration status (native=1)	-9.033*** (0.646)
Time spent on internet (TIMEINT)	-0.275*** (0.016)
Resources at school (SCMATEDU)	9.469*** (0.646)
Teacher-student ratio (STRATIO)	8.251*** (0.179)
School type (public=1)	39.511*** (3.801)
Computer – student ratio (RATCOMP)	8.251*** (0.179)
Student-Teacher mathematics ratio (SMRATIO)	-0.312*** (0.014)
Treatment variable as use of ICT (yes=1)	6.499*** (1.224)
No. obs	21,520
R-squared	0.24

Source: OECD-PISA 2012 data
Note: Result of linear regression with mathematic test score as the dependent variable. Cells show the marginal effects evaluated at the means of all explanatory variables. Robust standard errors are shown in brackets below. Superscripts ***, **, * denote statistical significance at the 1, 5 and 10 per cent levels respectively.