ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

THE EFFECT OF INSTRUCTION TIME ON STUDENT OUTCOME BY USING FIRST DIFFERENCE APPROACH WITH TIMSS DATA

M.Sc. THESIS

Eylül Berke ÖZGEN

Department of Economics

Economics M.Sc Programme

FEBRUARY 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

THE EFFECT OF INSTRUCTION TIME ON STUDENT OUTCOME BY USING FIRST DIFFERENCE APPROACH WITH TIMSS DATA

M.Sc. THESIS

Eylül Berke ÖZGEN

(412201015)

Department of Economics

Economics M.Sc Programme

Thesis Advisor: Asst. Prof. Dr. Aysun HIZIROGLU AYGÜN

FEBRUARY 2023

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

DERS SÜRESİNİN ÖĞRENCİ BAŞARISINA ETKİSİNİN BİRİNCİ FARKLAR YAKLAŞIMINA DAYALI OLARAK VE TIMSS VERİLERİ KULLANARAK ELDE EDİLMESİ

YÜKSEK LİSANS TEZİ

Eylül Berke ÖZGEN

(412201015)

İktisat Anabilim Dalı

İktisat Yüksek Lisans Programı

Tez Danışmanı: Dr.Öğr.Üyesi Aysun Hızıroğlu AYGÜN

ŞUBAT 2023

Eylül Berke Özgen, a M.Sc. student of ITU Graduate School, student ID 412201015, successfully defended the thesis/dissertation entitled ―The Effect of Instruction Time On Student Achievement By Using First Difference Approach With TIMSS Data, which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor: **Asst. Prof. Dr. Aysun HIZIROGLU AYGÜN** …….
Istanbul Technical University

Jury Members: **Asst. Prof. Dr. Shourjo CHAKRAVORTY** …….
Istanbul Technical University

Assoc. Prof. Dr. Abdullah TİRGİL …….
Ankara Yildirim Beyazid University

Date of Submission **: 25 January 2023**

Date of Defense **: 15 February 2023**

To the kids all over the world,

FOREWORD

I wanted to research in education which I believe the only cure of ignorence. With this thesis, my main concern is to add value to the society by suggesting a policy on education. I believe that the outcomes presented in this thesis will help the policy makers and education authorities for allocating appropriate instruction time at primary schools. Firstly, I thank God for letting me to finish this research. I could not end this work if he did not support me during all the challanges I faced. I thank my distinguished supervisor, Asst. Prof. Dr. Aysun HIZIROĞLU AYGÜN for her valuable guidance on this work to produce a good quality research paper. Finally, I thank my lovely wife and all of my family for their love and support.

March 2023 ..... Eylül Berke ÖZGEN

TABLE OF CONTENTS

Page

ABBREVIATIONS

FD : First Difference
FE :Fixed Effects
OECD :Organization for Economic Co-operation and Development
PIRLS :Progress in International Reading Literacy Study
PISA :Program for International Student Assessment
PV :Plausible Value
SD :Standard Deviation
TIMSS :Trends in International Mathematics and Science Study

LIST OF TABLES

Page

LIST OF FIGURES

Page

THE EFFECT OF INSTRUCTION TIME ON STUDENT OUTCOME BY USING FIRST DIFFERENCE APPROACH WITH TIMSS DATA

SUMMARY

Human capital is one of the key elements of economic development. Upon the importance of human capital, some detailed assessments on student achievement enable researchers to measure subject-specific abilities of students over countries. From this perspective, the International Association for the Evaluation of Educational Achievement (IEA) conducts detailed large-scale assessments on 8th and 4th grade students, named TIMSS, over countries. While some countries' pupils and schools have high scores, the rest obviously falls behind. These academic achievement gaps of pupils between countries and within countries have different and reasonable explanations in the overall literature. Our study's purpose is to enlighten one of the possible reasons for score variation and propose policy advice based on empirical outcomes.

Through this study, instructional time has been taken as one of the significant inputs in the education production function by including other control variables which were extracted from TIMSS 2019 dataset. Four different educational systems and countries were investigated and to avoid foregoing biases and obtain a causal relationship, first differences was used in a within-pupil setting. The focused dataset includes over 19,000 observations of pupils from four countries namely Canada, Azerbaijan, Slovak Republic, and Spain. The overall results which are based on the whole dataset show a significant and positive point estimate of the effect of instructional time. In this setting, the effect of mathematics instruction time on mathematics score was estimated as 0.050 point increase per minute/week and the effect of science instruction time on science score was estimated as 0.094 point increase per minute/week. In the country-specific analysis, three of four countries showed insignificant responses to the explanatory variable. Upon that outcome, in the second phase, only Canada which showed a significant and positive response to instructional time has been studied. In Canada, the effect of mathematics instruction time on mathematics scores was 0.061 per minute/week and the effect of science instruction time on science scores was 0.133 per minute/week. Afterwards, heterogeneity tests on gender and immigration status were examined. Outcomes showed that the group of girls and immigrants were affected more by instructional time additions than the boys and natives. Then, to check the correctness of the identification strategy, one placebo test was employed by using a uniformly random number generation method to produce the counterfeit instruction times and reallocate them with the original ones. The test resulted in insignificant outcomes in 98 out of 100 random variations in regressions as expected, meaning that significance of instructional time effect was not obtained based on chance or coincidence.

Besides, each country’s educational system was investigated as an interpretation by focusing on several key metrics including educational autonomy, government

expenditure on education, teacher-pupil ratio, teacher certification, academic output in the educational research, curriculum comprehensiveness and teacher salary. Within the boundaries of dataset, as a quantitative approach, logit and probit regressions were run to obtain the explanation of productivity differences among educational systems. Correspondingly, adequate time on each of the topics during class hours was found foremost for instructional time productivity of achievement while the teacher-pupil ratio does not have any effect. As an inference, approaching a self-deterministic structure of governance of schools and increased research outcomes for education sciences in the academic area were found to be other possible important metrics for instructional time productivity. Overall, the importance of this study comes from contributions to the literature by complementing the previous findings in several ways. Firstly, the response to the increased instructional time was investigated in four diverse countries and educational systems by quantitative methods including General Linear Regression, First Difference regression and Logit and Probit regressions. Secondly, several variables for productivity of achievement to instructional time were proposed. Thirdly, detailed research for heterogeneous effects on the productive country, Canada, was accomplished. Moreover, the law of diminishing returns for instructional time was examined and possible evidence for an optimum time of instruction was presented. Fourthly, a placebo test using random number generation to check the robustness of the significance of the instructional time effect was applied. In conclusion, additional instructional time may not always result always with increased student achievement. Rather, the subject country should be investigated carefully by considering its status quo education system characteristics (autonomy, curriculum intensity on the given time, demographics and research on education). On the other hand, as a specific analysis for the productive country Canada, science and mathematics instructional time effects are significantly positive for all pupils but higher for girls and immigrants. Also, science instruction's effect is higher than mathematics in all subgroups. In the literature, a considerable number of studies focus on the same subject. However, as student achievement scores are the final products of piles of diverse and complex inputs, the results of these studies may include biases. That includes selection bias when pupils are sent to schools/classes depending on their ability on a specific subject. Moreover, as the education production function includes family, student, school, class, teacher and even environment-based inputs, empirical outcomes from other studies may suffer from omitted variable bias and measurement errors.

DERS SÜRESİNİN ÖĞRENCİ BAŞARISINA ETKİSİNİN BİRİNCİ FARKLAR YAKLAŞIMINA DAYALI OLARAK VE TIMSS VERİLERİ KULLANARAK ELDE EDİLMESİ

ÖZET

Beşeri sermaye, ekonomik kalkınmanın temel unsurlarından biridir. Beşeri sermayenin önemi üzerine, araştırmacılar ülke bazında matematik ve fen gibi konuya özgü bilgi ve yeteneklerin ölçülmesini sağlayan ayrıntılı çalışmalar yürütmektedir. Bu bağlamda, International Association for the Evaluation of Educational Achievement (IEA), TIMSS adlı 8. ve 4. sınıf öğrencileri için birçok ülkeyi barındıran ayrıntılı ve geniş çaplı değerlendirmeler yapmaktadır. Bu değerlendirmeler sonucunda, bazı ülkelerin öğrencilerinin ve okullarının puanları yüksekken, bazılarının ise oldukça düşüktür. Ülkeler arasındaki ve hatta ülke içindeki akademik başarı farkları, genel literatürde farklı ve geniş kapsamlı açıklamalara sahiptir. Bu çalışmadaki amaç, puan değişiminin olası nedenlerinden birini aydınlatmak ve ampirik sonuçlara dayalı olarak politika tavsiyeleri sunmaktır.

Bu çalışma aracılığıyla, TIMSS 2019 çalışmasında çeşitli anketlerden elde edilen kontrol değişkenleri de dahil edilerek ders süresi, eğitim üretim fonksiyonunda önemli girdilerden biri olarak ele alınmıştır. Dört farklı eğitim sistemi ve ülke incelenmiştir. Asıl sebep sonuç etkisinin anlaşılmasının önüne geçebilecek hatalardan kaçınmak ve safi bir nedensellik ilişkisi elde etmek için birinci farklar ekonometrik yöntemi ana nicel yöntem olarak kullanılmıştır. Odaklanılan veri seti, Kanada, Azerbaycan, Slovakya Cumhuriyeti ve İspanya olmak üzere dört ülkeden 19.000'den fazla öğrenciyi içermektedir. Tüm veri setine dayanan genel sonuçlar, ders süresinin etkisinin anlamlı ve pozitif olduğunu göstermektedir. Bu model üzerinde matematik öğretim süresinin matematik puanı üzerindeki etkisi, ders süresinin haftalık bir dakikalık artışının 0,050 puan artışıyla sonuçlanacağı şeklinde tahmin edilmiştir. Fen öğretimi süresinin fen puanına etkisi ders süresinin haftalık bir dakikalık artışının 0,094 puan artışıyla sonuçlanacağı şeklinde tahmin edilmiştir. Ülkeye özgü analizlerde ise, dört ülkeden üçü ders süresi değişkenine anamlı yanıt vermemiştir. Bu sonuç üzerine, ikinci aşamada sadece öğretim süresine anlamlı ve olumlu yanıt veren ülke, Kanada incelenmiştir. Kanada'da, matematik ders süresinin matematik puanı üzerindeki etkisi haftada bir dakika artışında 0,061 puan ve fen öğretimi süresinin fen puanı üzerindeki etkisi haftada bir dakika artışında 0,133 puan olarak tahmin edimiştir. Ardından cinsiyet farkılığı ve göçmenlik durumuna bağlı olan heterojen durumlar incelenmiştir. Sonuçlar, eğitim süresi eklemelerinden kız ve göçmen grubunun erkeklerden daha fazla etkilendiğini göstermiştir. Daha sonra, sonuçların sağlamlığını araştırmak için, sahte ders sürelerini üreterek ve bunları orijinal olanlar ile değiştirerek, rasgele sayı üretme yöntemi ile placebo testi uygulanmıştır. Test, beklendiği gibi, regresyonlardaki 100 rastgele varyasyondan 98'inin anlamsız olarak bulunması ile sonuçlanmıştır. Bu sonuç ile belirtilmektedir ki çalışmada elde edilen sonuçlar şansa veya tesadüflere dayalı olarak elde edilmemiş, anlamlı bir etkinin var olduğu delillendirilmiştir.

Bunlara ek olarak, her ülkenin eğitim sistemi, eğitim özerkliği, eğitime yapılan devlet harcamaları, öğretmen sayısının öğrenci sayısına oranı, öğretmen sertifikasyonu, eğitim alanındaki akademik çıktı sayısı, müfredatın kapsamlılığı ve öğretmen maaşı gibi birkaç temel ölçüte odaklanarak incelenmiştir. Logit ve probit regresyonları kullanılarak, ders saatlerinde konuların her biri için yeterli zaman ayrıması, başarının öğretim süresi verimliliği önemli bir unsur olarak bulunurken, öğretmen öğrenci oranının bu verimliliğe herhangi bir etkisi saptanmamıştır. Ayrıca, okulların otonom bir yapıya yaklaşması ve daha kaliteli eğitim için akademik alandaki araştırmaların artması diğer olası önemli belirteçler olarak yorumlanmıştır.

Bu çalışmanın önemi literatüre sağladığı özgün katkılardan kaynaklanmaktadır. İlk olarak, artan öğretim süresine verilen yanıt, Genel Doğrusal Regresyon, Birinci Fark regresyonu ve Logit ve Probit regresyonları dahil olmak üzere hem nitel hem de nicel yöntemlerle dört farklı ülke ve eğitim sistemi incelenmiştir. İkinci olarak, yapılan analizler temel alınarak, öğretim süresinin başarıya verimliliği için birkaç farklı değişken ayırt edilmiş ve önerilmiştir. Üçüncüsü, Kanada özelinde heterojen etkiler için ayrıntılı araştırma gerçekleştirilmiştir. Ardından öğretim süresi için azalan verimler yasası incelenmiş ve optimum öğretim süresi için kanıtlar sunulmuştur. Dördüncüsü, öğretim süresi etkisinin öneminin sağlamlığını kontrol etmek için rasgele sayı üretimi yöntemi kullanan bir plasebo testi uygulanmıştır. Sonuç olarak, belirli bir ülkede ders süresi arttırımları, her zaman yüksek başarı getirisi ile sonuçlanmamaktadır. Aksine, politika uygulanmadan önce ilgili eğitim sistemi dikkatlice araştırılmalıdır. Öte yandan, Kanada'ya özgü bir analiz olarak, fen ve matematik derslerinde, ders süresinin marjinal etkilesi tüm öğrenciler için anlamlı şekide pozitif, kızlar ve göçmenler için daha yüksek olarak saptanmıştır. Ayrıca fen dersi süresinin marjinal etkisi tüm alt gruplarda matematiğe göre daha yüksek olarak hesaplanmıştır.

Literatürde aynı konuya odaklanan önemli sayıda çalışma bulunmaktadır. Bununla birlikte, öğrenci başarı puanları, çeşitli karmaşık girdi yığınlarının nihai ürünleri olduğundan, bu çalışmaların sonuçları asıl etkiden sapmalar içerebilir. Örneğin, öğrencilerin belirli bir alandaki yeteneklerine bağlı olarak okullara/sınıflara yerleştirilmesi seçim yanlılığı durumuna sebep olabilmektedir. Ayrıca, eğitim fonksiyonu aile, öğrenci, okul, sınıf, öğretmen ve hatta çevreye dayalı girdileri içerdiğinden, ampirik sonuçlar ihmal edilen değişken yanlılığından ve ölçüm hatalarından dolayı yanıtıcı olmaya açık olabilmektedir.

1. INTRODUCTION

Academic achievement gaps of pupils between countries that are measured by large-scale assessments such as TIMSS, PIRLS and PISA may have different reasons and one of them is seen as the instructional time in the related literature. There are governments, international organizations, and other authorities that value instructional time as an input that enhances learning.[1] This is highly likely true when these are the cases where increasing instructional time enables teachers to deep into topics and have more question-and-answer times. Moreover, with the easy-to-adapt mechanism of instructional time additions as no more infrastructure investment is required (except for the need of supplying more energy and maintenance bills occurring from the over-use of current infrastructure and also additional payments for teachers), it could be an attractive expedient to be applied from the side of policymakers. Besides, while having these facts, there is an excessive diversity of instructional times between nations. For example, according to the Education at Glance 2019 Report, the OECD average of compulsory primary instructional time is 4568 hours, annually. While the range is between 2769 hours (the compulsory instruction time for Hungary) and 7000 hours (the compulsory instruction time for Australia)[2].

Furthermore, there is a considerable number of studies on the purpose of investigating instructional time's effect on student scores by having data from large-scale assessments like TIMSS and PISA. However, the findings show mixed results which makes it harder to have a consensus.[3] In addition, some potential biases may be the case. Namely, selection bias occurs when parents choose schools for their kids

[1]Among others, a few examples: Firstly, On the NYC Elementary School Academic Policy Guide, September 2022 minimum 900 hours instruction for K-6 grade is declared. Secondly, UK Aid's Sierra Leone Education Policy recommendation 2020, clearly underlines "Increasing Instructional Time to Improve Learning Achievement".

[2] In the report of OECD (2019) Education at a Glance 2019, the compulsory instruction times are presented for OECD countries at the Figure D1.1.

[3] While Long (2014) found no effect of reading instruction on reading achievement, only a small effect of mathematics instruction on mathematics achievement by using PISA 2000, Lee and Barro (2001) found no significant effect of school year duration and Shaw et al. (2013) found no effect of course duration on learning outcomes, contrarily Lavy(2015) and Cattaneo et al. (2016) found positive effect of instructional time by using PISA 2006 and PISA 2009 data, respectively. More other studies exist with different findings.

providing higher instruction time or when pupils are located in classes providing different instructional time programs by school governance based on pupils' ability. Other sources of possible bias are unobserved characteristics. While data collection process, surveys or large-scale assessments may not provide enough and/or suitable information about subject under research. Appropriately, the exact causation mechanism behind the relationship between instructional time and achievement based on test scores is not straightforward to detect. Besides these issues, instructional time includes different spectrums as one of them, is "actual lesson time" which is the actual instructional time after removing closure of schools due to official reasons, absenteeism or late attendance of students or organizational breaks within the school year; and as another spectrum, "engaged learning time" which is the actual learning time that student is engaged and pay attention to the instruction.[4] Because we cannot observe the efficiency of classes, all the studies are potentially biased as they all, and this study too, take one component by omitting its in-depth structure. However, it is obvious that an empirical consistency exists among studies which use diverse samples and research settings. By reclining this fact, notwithstanding the arduousness of the measurement of the causal effect, a part of this study contributes to this "empirical regularity"[5] by sharing parallel and also detailed results in its scope.

Through this study, instructional time was treated as one of the significant inputs in the education production function. In the first phase of the study, four different educational systems were chosen and to extract causal relationships well-known econometrics method, the first differences, and within-pupil setting were used. While the whole dataset which includes over 19,000 observations of pupils showed a significant and positive point estimate of the effect of instructional time. In the country-specific three out of four countries, Azerbaijan, Slovak Republic and Spain showed no significant response to the explanatory variable was measured. After obtaining outcomes and differences become obvious among country results, an interpretation and a quantitative analysis based on logit and probit regressions were implemented to explain country-based variation. Then in the second phase, for one of the countries, Canada which was productive to the instructional time additions, detailed investigation was accomplished by examining heterogeneity and diminishing

[4] In the study of "What's All the Fuss About Instructional Time?", 1990, Berliner explains the spectrums of instruction time by the page 2.

5 The propositional phrase is mentioned in Lavy (2015)

returns of time. Results similar to the literature are found.[6] Also similar to Lavy (2015), the effect of instructional time was found to be larger for girls and immigrants. Then, one placebo test was undertaken by using a uniformly random number generation method to produce and replace the lookalike instruction times.

This study contributes to the literature in four ways as follows: First, as a unique contribution to our knowledge, the productivity of achievement to instructional time additions over the countries with four different education systems is analyzed and detailed education system analysis for these countries is presented. Secondly, based on the results in the first phase, detailed research for heterogeneous effects on the instructional time-productive country, Canada, is accomplished. Thirdly, the law of diminishing returns in the duration of instructional time is scrutinized and presented as possible evidence for an optimum time of instruction. Fourthly, a placebo test is presented using random number generation to be certain about the significance of the instructional time variable. At the end of these, it is determined that adequate time on each of the topics during class hours is foremost for instructional time productivity of achievement while the teacher-pupil ratio has not any effect. Also, approaching a self-deterministic structure of educational governance of schools and academic research for quality of education are found as other possible pillars of instructional time productivity of achievement. In the instructional time-productive country, Canada, within-pupil FD regression offers a much lower effect than the general regression, but the result creates a marginal effect of 0.061 points per minute/week and 0.133 points per minute/week for mathematics and science scores, respectively, which contributes the empirical consistency in this research subject. Also, the outcomes of this study support previous findings by sharing higher marginal effects of instructional time on achievement for girls and immigrants.

The rest of the article is organized as follows: Section 2 presents the existing literature. Section 3 describes and presents the dataset and the sampling design of the TIMSS 2019. Section 4 states the empirical strategy by highlighting three significant linear equations in which instructional time is the explanatory variable. This section also describes the potential sources of biases and the assumptions that have been taken.

6 Lavy (2015) found 5.8% SD change in case of an hour increase in instruction by taking the study's 2006 PISA sample into consideration by using his FE Model setting as a general result for mathematics, science, and language subjects. As 1 SD equals to 84.4 for the dataset, results imply 4.89 points increase. Cattaneo et al. (2016) found 5.9% change with similar setting but using study's Swiss students 2009 PISA sample. As 1 SD equals to 85.3 for the dataset, results imply 5.03 points increase. As implicated in Results section, this paper's point estimates are very close to the mentioned paper supporting the empirical regularity.

Section 5 and 6 presents a detailed analysis of the educational systems of the studied countries and followingly, focuses on Canada. Section 7 concludes.

2. LITERATURE REVIEW

As instructional time is a source that quickly and easily be adapted to the educational system, its investigation on educational output is valuable for policy makers. Instructional time's effect on educational achievement has been studied by several papers through the history because of that and other relatedly varying reasons.

In addition to the papers that directly investigates the effect, a few papers presented correlations or interpretations over the subject. Dağlı (2019), as a recent work, has investigated and putted together related literature and provided sufficient overview on the topic. OECD (2011a) Report which uses the PISA 2006 survey, concludes that a positive correlation exists between reading, mathematics and science scores and regular school day instruction time on 15 years old student over participant OECD countries. The mean instructional time during the lessons helped to explain between-country variance by 25% in reading and mathematics success according to the report. Farbman et al. (2014) which bases on the US education system, pointed out that more and high-quality learning time with specialized curriculum affects disadvantaged students at most.

The papers which are based on a causal relationship can be listed in the following passages. Lavy (2015) published one of the most citable results on PISA 2006 survey which is based on 15 years old student's assessment over participant OECD countries. The research was built on estimating fixed-effect within-student regression for instruction time's effect on reading, mathematics, and science, with controls that are time-invariant. 22 OECD countries took part in the study with more than 400,000 observations. He pointed out that one hour addition to the weekly plan of lessons in reading, mathematics, and science classes affected positively and increased the PISA test scores by around 6% of a SD (84.4). In heterogenous analysis, he found that the effect was lower for boys than for girls. Also, the effect was lower for the pupils whose parents had high education. Moreover, his results pointed that the effect of one-hour additional instruction time was also lower for the developing countries rather than Eastern European countries. Eide and Showalter (1998) studied the effect of the

duration of a school year in USA on success by using the data from high school sophomores. Finally they concluded with the insignificant results. Fryer (2011) published more comprehensive study as focusing on some educational reforms' effect on the student success in Houston USA. He used student observations and the grades varying from third through eleventh grades. He argued that despite of a lot of reforms over the country, USA has had the same place in the large-scale assessments which measures academic achievement of student among other countries. On the other hand, as the paper presented, the application of a mix of comprehensive reforms including increased learning time/more time on the task to the schools, which was the case for Houston charter schools, results with higher achievement scores on math and reading. Rivkin and Schiman (2015) used PISA 2009 data with over 500,000 observations including only 9th and 10th grade students over 72 countries. By using school-by-grade-by-subject level FE, they found that achievement scores can increase with additional instructional hours. However, this increase was higher for the students who were top half of the class in academic achievement ranking. They emphasized that additional time possibly let teacher cover the topic more in detail, to tailor the topics for individual needs of specific students and to answer the questions more. They also found that additional hours can give benefits to the student, but the amount of the addition and classroom factors are other inputs of overall gain. Rivkin and Schiman (2015) reported 0.051 as marginal coefficient of score for each additional minute. As nearly same results can be seen on my study[7] Long (2014) applied a three-level approach with country, school and student attributes to two datasets: PISA 2000 and PISA 2006, both uses 15 years old students over OECD countries. He found no effect of reading instruction on reading achievement, a small effect of mathematics instruction on mathematics achievement, and no effect of mathematics instruction to the achievement at the country or school levels in PISA 2000 study (between countries and between schools). In contrast, on the same paper but with PISA 2006 dataset, an additional hour of mathematics or reading instruction raised mathematics and reading test scores by half of a SD at country level. Also, students with a low socioeconomic background benefitted from the additional hour of instruction more than those with a higher socioeconomic background. The latter is contrary to Lavy (2015)'s findings as

[7] In the study of Rivkin and Schiman (2015), this is only valid for the lectures which has 240-279 minutes per week as instructional time and only for 2 subjects(math and language arts).

the latter found a better effect on the student whose parents had higher educational attainment. Long (2014) concludes with raising need for further studies in order to be certain about increased instructional time's effect on achievement. Jensen (2013) presented variation in classroom hours induced by a reform that changed instructional time by 2.2-3.3% in Denmark. He found that one-hour increase per year in mathematics instruction increased scores on mathematics tests by 0.21% of standard deviation at 9th grade. Blank (2013) used National Assessment of Educational Progress (NAEP) 2009 data which includes 46 states and 2000 students per state on grade 4 students for analyzing instructional time effect on science scores. His results demonstrated that instructional time of science lectures has positive effect on grade 4 students. In detail, one state with 120 minutes per week science instruction has 150 points on average while a state with 210 minutes per week science instruction has 154 points. The 4-point difference has been found statistically significant. Also, they reported that the students receiving high amount of science instruction (4 hours per week in the study) were 12 points better than the ones receiving the lowest amount of instruction (1 hour per week in the study). This finding is also very similar as 12 points difference in 180 minutes makes each minute 0.066 point worth which is very close to my conclusions mentioned firstly in the Introduction section of this paper. Cattaneo et al. (2016) made a replication of Lavy (2015)'s study with Swiss 15 years old students by using PISA 2009 data. They also found similar results with their reference article: one additional hour increase in instruction has resulted with around 6% standard deviation of scores. They checked for several heterogeneity issues as well. Their findings showed that high ability students benefitted more as 8% SD increase found while in low ability tracks this effect was only 3.8% of SD. Woessmann (2003) used TIMSS database with a cross-country study including 39 countries and 7th and 8th grade students. He presented positive effect of instructional time on mathematics scores by using international. Baker et al. (2004) used three international datasets (PISA 2000, TIMSS 1999, CIVICS 1999) to analyze the causation between instruction time and student achievement scores in huge country set and reported no significant effect. They concluded with only little evidence of existence of marginal effect to the scores.

Moreover, Pischke (2007) used historical database dated 1961–2 to 1972–3 from Germany on primary and secondary level students. He checked the impact of duration

of school year on future wages by using SUR (seemingly unrelated regression) and Student FE estimators but could not find any statistically significant estimates. However, Pischke (2007) found that short school years was significantly related with grade repetition and lower track choice for future education of student. Similar study has been done by Lee and Barro (2001) used 58 countries data on different age groups for research mathematics, reading and science subjects on learning achievement. Resulted with no significant effect of school year. Shaw et al. (2013) investigated the course length effect, rather than the school length, on the undergrade level students' achievement in USA. They compared whether there was a difference in student achievement between 8-weeks and 16-weeks length online courses about abnormal psychology. They concluded that there was no significant difference between these two settings.

Also, Tiedt et al. (2021) found no significant difference on achievements for their experiment on nurse graduate education on the 8 weeks or 16-weeks courses in USA. As a recent study on same issue, Satoh et al. (2022) has different setting. They designed a lab-based lectures on several themes(subjects) of learning for second grade medical students in 2018 and followingly 2019 in Japan, Saitama Medical University. In 2018, each theme was given for one day while in 2019 each theme was given for two days duration, but the educational input weight was same. They applied 5-point Likert chart survey to the second-year undergraduate medical students for several learning outcomes including understanding. Finally, they have seen no difference between two years' understanding results. As an example of out of class instructional time, the amount of time spent on after-class homeworks has also been related with positive effect on student achievement(Rønning,2010).

3. DATA

The International Association for the Evaluation of Educational Achievement (IEA) organizes The Trends in International Mathematics and Science Study (TIMSS) in every four years. TIMSS shares timely information on the education system of participant countries, and its 4th and 8th grade students' capability on mathematics and science subjects as well as background information of students such as parental support at home, sense of school belongings, enjoyment during lectures and more . The first TIMSS was held in 1995, however this paper uses the data from TIMSS 2019 which is the 7th cycle and participated by 64 countries and several provincial participants. To ensure statistical correctness and comparability between countries, the IEA follows predefined requirements during the data collection process. The target population for 4th grade students, who are the main target group within this study, are all students enrolled in the 4th grade and it is underlined that the mean age is 9.5 years old at the test time.[8] Besides, the target population for teachers are all mathematics and science teachers who are the teachers of sampled students. As it is not possible to test all population, a sampling mechanism is being used by the TIMSS officials.

Before proceeding the phase of school section, the "National Research Coordinator" who is a country-officer and appointed for coordinating TIMSS procedures in her country, constructs a "sample frame" that consists of all schools in the country for target grade properly. Then, based on this frame, two staged random selection process being held by TIMSS officers. In these stages, schools are randomly selected and followingly, classes are randomly selected. Firstly, based on probability-proportionate-to-size (PPS) approach, some individual schools are selected. PPS is an approach that creates a school-specific probability score which is proportional to related school's entire target grade population. This implies that more observation in the target grade(4th and/or 8th grade students) results more chance for being selected for this year's TIMSS examination. However, participation is optional. If selected school do not want to attend the study, TIMSS statisticians appoint a substitute school

[8] Related statistics can be found detailly at Methods and Procedures: TIMSS 2019 Technical Report, Chapter 3.

which is in the same neighborhood. Also, some exclusion decisions of the schools may be the case while selection, based on the detailed criterions decided by the TIMSS.[9] In the second stage, classes are select by TIMSS & PIRLS sampling software that uses a standard algorithm for all selection processes. It is possible to select one or more classes per school. Overall, the sample is formed by a random selection of all schools and the classes in related country.[10]

The achievement scores for the subjects in the studies such as TIMSS and PISA and other possible large-scale assessments are the plausible values. These variables are counted as measures of students' ability on given tests. There are 5 plausible values reported per subject in TIMSS'19. Basically, PVs are used to estimate the subject-based ability of whole population of selected country. Individual PVs can be seen as the estimated values based on students' responses to items and background information. Each PV is randomly drawn from a distribution which is based on an estimation and being considered as a representative value for all students who have alike pattern of responses and similar backgrounds. Differences between PVs are caused by the degree of errors in the distribution of scores.[11]

The data on this study was retrieved from files with "-M7" suffix extension on the TIMSS 2019 database which is based on 7th cycle of TIMSS and consist of eTIMSS (TIMSS questionary on computer) and paper-based TIMSS (TIMSS questionary on paper). Other data types, bridge and eTIMSS with PSI (Problem Solving and Inquiry), are not used in this study.

This study firstly uses TIMSS data from four different countries which are Canada, Azerbaijan, Slovak Republic, and Spain. Then, specifically address the data for Canada. These four countries have different historical backgrounds. For example, while Azerbaijan and Slovak Republic is newly constructed republics after the dispersion of Soviet Union and Czechoslovakia, Canada mostly formed in 19th century with a huge influence from European settlers, Spain has its foundation roots from 15th century.[12] The diversity in the histories and also, the geographies of these countries

[9] On the page of National Center for Education Statistics (NCES), https://nces.ed.gov/timss/timss15technotes_intlreqs.asp#f1) detailed information can be seen.
10 Procedures are underlined in the book of Methods and Procedures: TIMSS 2019 Technical Report, Chapter 3
11 For more detail info about PV's used in TIMSS, Methods and Procedures: TIMSS 2019 Technical Report, TIMSS and PIRLS Achievement Scaling Methodology-TP11 are the main resources.
[12] For more detailed information about Spanish History please visit the book of Britannica, T. Editors of Encyclopaedia. "Catholic Monarchs." Encyclopedia Britannica, December 2, 2013

led them form different systems of governance in politics and as well as in education. These four countries were used in this study to investigate productivity of "instructional time" across different educational systems, to explain the reason behind the formation of productivity of instructional time, and to have detail research for a country which is productive in terms of instructional time.

As seen at Table 3.1, Azerbaijan has the highest score in mathematics while has the lowest score for science, 528.9 and 445.0, respectively. Canada is around in close to the average for science, with 516.3 point. Slovakia has the highest score for science and above the average for mathematics, 525.3 and 512.4, respectively. Spain is also close to average for both subjects, 513.8 point for science and 504.3 point for mathematics. For lecture durations, Canada and Spain are above the average for both lectures while Slovakia and Azerbaijan are below the average. As, on average, Canada has 311.4 minutes per week for mathematics and 137.7 minutes per week for science; Spain has 273.0 minutes per week for mathematics and 177.1 minutes per week for science; Azerbaijan has 215.2 minutes per week for mathematics and 90.3 minutes per week for science; Slovak Republic has 197.3 minutes per week for mathematics and 94.8 minutes per week for science. Moreover, the standard deviation is the highest in Canada for both lectures, followed by Spain. In Azerbaijan, possibly due to the central educational system, both lecture durations are not dispersed from the mean.

Table 3. 1: Data Description.

	Whole Sample		Azerbaijan		Canada		Slovakia		Spain	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Mathematics Score	507.6	76.9	528.9	82.0	498.2	75.9	512.4	72.6	504.3	72.9
Science Score	502.7	81.3	445.0	93.6	516.3	71.0	525.3	77.0	513.8	67.7
Mathematics Lecture Duration	268.2	78.2	215.2	27.0	311.4	87.0	197.3	32.2	273.0	47.3
Science Lecture Duration	131.7	68.0	90.3	6.9	136.7	75.0	94.8	20.6	177.1	68.1
No. of Observations	19704	19704	3864	3864	8516	8516	2608	2608	4716	4716

Moreover, at Table 3.2, all the background data which are used as control variables are with their details. The names of variables, which are used in this paper, the types of variables, the names of variables in the TIMSS 2019 database and the data about their extractions have shared for clarity and reproducibility of the work.

Table 3. 2: Variables and Descriptions.

Variable	**Type of Variable**	**Name of the variable in TIMSS'19 Database**	**Detail of Variable**
Math Achievement Score	Continuous	ASMMAT01 to ASMMAT05	Mathematics achievement scores are the dependent variables in the study. Each of the scores is a plausible value which described in the "Data" section. (National Center for Education Statistics)
Science Achievement Score	Continuous	ASSSCI01 to ASSSCI05	Science achievement scores are the dependent variables in the study. Each of the scores is a plausible value which described in the "Data" section. (National Center for Education Statistics)[13]
Mathematics Instructional Time	Continuous	ATBM01	Math Instructional Time variable is obtained from the question of "In a typical week, how much time do you spend teaching mathematics to the students in this class? (In minutes)" in the TIMSS'19 teachers' survey.
Science Instructional Time	Continuous	ATBS01B	Science Instructional Time variable is obtained from the question of "Please estimate the time that you spend on science topics with students in this class (in minutes per week)" in the TIMSS'19 science teachers' survey.
Born in Country	Categorical	ASBG07	Variable is obtained from students' survey. Two answers as categories are allowed: "1: Yes" and "2: No".
Gender	Categorical	ASBG01	Variable is obtained from students' survey. Two answers as categories allowed: "1: Girl" and "2: Boy".
Age	Continuous	ASDAGE	Variable is obtained from students' survey.
Sense of School Belongings of Pupil	Continuous/Scale	ASBGSSB	Variable is obtained from students' survey as a scale named "Students Sense of School Belonging/SCL" which is extracted from 5 statements' evaluation by the pupil. Then, a scale is formed as a continuous number by TIMSS'19. These 5 statements that are being evaluated from pupil by marking 4 options between "Agree a lot" and "Disagree a lot" options are listed as: "I like being in school" , " I feel safe when I am at school" , "I feel like I belong at this school", "Teachers at my school are fair to me", " I am proud to go to this school"
Parental Support	Continuous/Scale	ASBGHRL	Variable is obtained from students' survey as a scale named "Home Resources for Learning/SCL" which is extracted from 5 questions' evaluation by the pupil and parent. Then, a scale is formed as a continuous number by TIMSS'19. These 5 questions that are being answered are listed as: "Number of books in the home (students)", "Number of children's books in the home (parents)", "Number of home study supports (students)", "Highest level of education of either parent (parents)", "Highest level of occupation of either parent (parents)" .

[13] National Center for Education Statistics shares detail information about weights and the procedure on the following page: (https://nces.ed.gov/timss/timss15technotes_weighting.asp

Table 3.2(Cont.): Variables and Descriptions.

Variable	Type of Variable	Name of the variable in TIMSS'19 Database	Detail of Variable
Like Learning Math	Continuous/Scale	ASBGSLM	Variable is obtained from students' survey as a scale named "Students Like Learning Mathematics/SCL" which is extracted from 9 statements' evaluation by the pupil. Then, a scale is formed as a continuous number by TIMSS'19. These 9 statements that should be evaluated includes “I enjoy learning mathematics", "Mathematics is boring", "I like to solve mathematics problems", " I wish I did not have to study mathematics" statements and 5 similar statements more.
Like Learning Science	Continuous/Scale	ASBGSLS	Variable is obtained from students' survey as a scale named "Students Like Learning Science/SCL" which is extracted from 9 statements' evaluation by the pupil. Then, a scale is formed as a continuous number by TIMSS'19. These 9 statements that should be evaluated includes “I enjoy learning science", "Science is boring", "I like to solve science problems", " I wish I did not have to study science" and 5 similar statements more.
Class Size	Continuous	ATBG10A	Variable is obtained from teachers' survey upon the question of "Number of students in the class?"
Education of Teacher	Categorical	ATBG04	Variable is obtained from teachers' survey upon the question of "Level of formal education completed?". Available 7 answers as categories are listed as : 1: Did not complete <Upper secondary education—ISCED Level 3>; 2: <Upper secondary education—ISCED Level 3> (If you have not completed <post-secondary or tertiary education>, you will; 3: <Post-secondary, non-tertiary education—ISCED Level 4>; 4: <Short-cycle tertiary education—ISCED Level 5>; 5: <Bachelor’s or equivalent level—ISCED Level 6>; 6: <Master’s or equivalent level—ISCED Level 7>; 7: <Doctor or equivalent level—ISCED Level 8>
Age of Teacher	Categorical	ATBG03	Variable is obtained from teachers' survey upon the question of "Age of teacher?". Available 6 answers as categories are listed as :1: Under 25; 2: 25–29; 3: 30–39; 4: 40–49; 5: 50–59; 6: 60 or more
Gender of Teacher	Categorical	ATBG02	Variable is obtained from teachers' survey. 2 answers as categories are allowed: 1: Female and 2: Male
Native Language of Overall Pupils	Categorical	ACBG04	Variable is obtained from principles' survey. 5 answers as categories are allowed: 1: More than 90%; 2: 76 to 90%; 3: 51 to 75%; 4: 26 to 50%; 5: 25% or less
Economic Disadvantage of Overall Pupils	Categorical	ACBG03A	Variable is obtained from principles' survey. 4 answers as categories are allowed: 1: 0 to 10%; 2: 11 to 25%; 3: 26 to 50%; 4: More than 50%
Country Dummy	Dummy	-	The variables are created as dummies for each country (Canada, Azerbaijan, Slovak Republic, Spain) by the author based on the default variable of "IDCOUNTRY" from TIMSS'19 Database for each pupil.

Table 3.2(Cont.): Variables and Descriptions.

Science Curriculum Fulness	Continuous (%)	ATDSLIF ATDSPHY ATDSEAR	Variable is created from averaging 3 variables from the Teacher Survey of TIMSS 2019 Dataset. These 3 variables are response of each teacher on their class. The variables are the ratio of the covered subtopics (Life Science, Physical Science and Earth Science) in science lectures in current year or the previous year. Then, overall coverage of topics is obtained.
Mathematics Curriculum Fulness	Continuous (%)	ATDMNUM ATDMGEO ATDMDAT	Variable is created from averaging 3 variables from Teacher Survey of the TIMSS 2019 Dataset. These 3 variables are response of each teacher on their class. The variables are the ratio of the covered subtopics (Data, Number, Measurement and Geometry) in mathematics lectures in current year or the previous year. Then, overall coverage of topics is obtained.

4. EMPIRICAL STRATEGY

This section consists of the empirical setting for the effect of instructional time on the 4th grade pupil achievement by considering a set of control variables that are presented in "Data" section. The first model to explain instruction duration's effect on educational output in mathematics and science can be proposed by following linear regression education functions:

$$Y_{ijkm} = F_i\ \beta_{xm} + Z_{jck}\,\beta_z + \ \beta_{mm}W_{jcm} + \ \beta_{sm}W_{jcs}\ + \beta_c D_c\ + \ \epsilon_1 \qquad (4.1)$$

$$Y_{ijks}\ = F_i\ \beta_{xs}\ +\ Z_{jck}\,\beta_z + \beta_{ms}\,W_{jcm}\ \ +\ \ \beta_{ss}W_{jcs}\ +\ \beta_c D_c\ +\ \epsilon_2 \qquad (4.2)$$

where:

$$\epsilon_1 = \mu_k + \ n_j + p_i + s_c + e_{ijkm} \qquad (4.1a)$$

$$\epsilon_2 = \mu_k + \ n_j + p_i + \ s_c + r_{ijks} \qquad (4.2a)$$

Y_{ijkm} and Y_{ijks} are the achievement outputs for student i in mathematics and science, respectively. F_i is a vector including family and student background information (age, gender, parental support, sense of school belongings, science lecture enjoyment, mathematics lecture enjoyment and immigrant status), Z_{jck} is another vector containing the information from school j (native language of overall pupils registered to the school, economics disadvantageous of overall pupils registered to the school, the size of related class in the school) and teacher k (education, gender, age). W_{jc} contains the information of instructional time in terms of minutes, m for mathematics and s for science lectures. As seen, to catch spillover effect of mathematics and science instruction on each one, both effects are put on the other's linear equation which can be seen in Rønning (2010). In this regard, β_{mm} in Equation (4.1) stands for mathematics instruction's effect on mathematics achievement, β_{ms} in Equation (4.2) stands for mathematics instruction's effect on science achievement. Similarly, β_{sm} in Equation (4.1) stands for science instruction's effect on mathematics achievement, β_{ss} in Equation (4.2) stands for science instruction's effect on science achievement. D_c is the country dummy as we have four countries in the first phase of the study. In both

equations, student background information's coefficients have indicators as "m" and "s" because the data includes subject based taste for classes which can be seen in detail in "Data" section.

The potential unobserved individual effects for teacher, school, student, and class are marked respectively with $\mu_k,\ n_j,\ p_i,\ s_c$. Following terms, e_{ijkm} and r_{ijkm}, are the idiosyncratic error terms for mathematics and science educational output equations. μ_k may be available for teacher based negative or positive inputs for example low motivation for the job as negative side or having high professional development as positive side. n_j and s_c may again exist if school/class has been managed poorly or extremely well. p_i captures several unobserved characteristics such as inner ability and motivation of pupil or physical wellbeing for education and so on.

Having current regression' estimates high-likely includes some biases whose sizes and directions are not easy to detect. One of the greatest potential sources of bias is unobserved individual characteristics mentioned above. These may have correlations with mathematics and science lessons that prevent us see exact causation. Secondly, teachers or the decision maker of the lecture durations may allocate more time for low ability classes to be sure about they also have enough instruction. If that is the case, high instruction times are correlated with with lower achievement scores which serves as downward bias. Similarly, a selection bias may occur if parents select schools or classes enabling high instruction times for the high ability or low ability pupils; these are the unobserved effects that should be prevented for the research.

Below, in Equation (4.3), the solution for potential biases is addressed:

$$\Delta Y = F_i\ \alpha_1 + W_{jcm}\,\alpha_{mat} + W_{jcs}\,\alpha_{sci} + u_{ijks} \tag{4.3}$$

where:

$$\alpha_1 = (\beta_{xm} - \beta_{xs}),\ \alpha_{mat} = (\beta_{mm} - \beta_{ms}),\ \alpha_{sci} = (\beta_{sm} - \beta_{ss}) \text{ and the } u_{ijkm} = (e_{ijkm} - v_{ijkm}). \tag{4.3a}$$

This equation built on the following assumption: the pupils who have higher interest or ability in mathematics (or science) are not allocated to the classes with higher instruction time on mathematics (or science) rather than the other one. If this is the case, the error term could be correlated with the dependent variable despite the control variables and the first difference modeling. But, summarized in the "Country Specific Education Systems" section, there is no evidence about subject based schools for 4th

graders in any of the countries studied in this paper. As one last issue, at the cost of reduced sample size, the sample is limited to students who have the same teacher for both lectures with the aim of eliminating the teacher-unobserved effect(after the elimination, the remaining sample size is about %47 of whole dataset). Lastly, the estimates should be interpreted with attention as the subject-specific inner ability of pupil and subject-specific teacher reliance are not excluded albeit we excluded general inner-ability for teacher and pupil.

For Equation (4.3), $\alpha_{mat} > 0$ precisely implies that a student with longer instruction time in mathematics is more successful in mathematics compared to science, in contrast to a student with shorter instruction time in mathematics. In other words, as the duration of the mathematics course increases, the student's score in mathematics diverges positively from science score. If it is negative, the lecture duration effect on mathematics achievement is negative. Furthermore, a negative α_{sci} coefficient indicates that a student with longer time in science lessons is more successful in science compared to mathematics, in contrast to a student with shorter time. In other words, as the science instruction duration increases, a student's science score diverges positively from her mathematics score. If α_{sci} is positive, lecture duration time's effect on science achievement is negative. In the "Results" section, findings are shared with correction of α_{sci} as multiplied by -1 to make it easy to interpret and compare with mathematics instruction time.

As TIMSS User Guide for the International Database (2019) explains, 5 plausible values were used to run regressions. PISA, TIMSS and some other international large-scale assessments employ plausible values to obtain student achievement which has different aspects while measuring as discussed in the "Data" chapter previously. The averaged coefficient estimations are calculated from the software(R) being able to calculate and report results as TIMSS'19 recommends. Standard errors are also calculated upon the sampling variation of coefficient estimates and by considering the variability occurring from imputation of 5 plausible values. SDs may be calculated with bootstrapping as seen in Rønning (2010) and possible other studies. For the sampling variability, the standard errors are clustered at class/teacher level and for the regression itself, student total weights ("totwgt" variable from the TIMSS'19 database) are used to comply with recommendations of the official TIMSS user guide.

5. GENERAL RESULTS

Table 5.1 and Table 5.2 present the effect of mathematics and science instructions on the subjects for 4th grades from estimated Equation (4.1) and Equation (4.2) by general linear regression, respectively. Each column has a different equation based on the existence of control variables but includes all the country dummies in the equation.

In the first column in Table 5.1 which does not include any control variables but country dummies, the point estimate of instruction duration of mathematics lessons is 0.057 and statistically significant at a 99% significance level. This means that 4th grade pupils who have 1 hour more math hour in their weekly curriculum score 3.42 points higher, on average, in the test. Besides, the spillover effect of science instruction is negative but not significant. The second column in Table 5.1 includes all background controls presented in the "Data" section while keeping country dummies. The point estimate of instruction duration is much higher, 0.082 at the 99% significance level. This means a stronger effect which indicates that 1 more hour of math class in the weekly curriculum results in 4.92 points higher. However, the science spillover effect is measured as negative 0.095 meaning that increasing science instruction duration by 5.7 points decrease in math score. On the other hand, these results may include biases as potential endogeneity issues are available which are discussed in "Empirical Results" section.

Table 5. 1: Effect of instruction on mean TIMSS math test.

	(1)	(2)
Mathematics Instruction	0.057**	0.082***
	(0.029)	(0.018)
Science Instruction	-0.048	-0.095***
	(0.036)	(0.019)
Country Dummies	x	x
All Controls		x
No. of Obs.	19704	11997

Reported numbers are based on weighted linear regressions over one subject (Math). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

In the first column in Table 5.2, the point estimate of instruction duration of science lessons is 0.001 and statistically insignificant. Besides, the spillover effect of mathematics instruction is also insignificant. The second column in Table 5.2, that includes all background controls presented in the "Data" section while keeping country dummies. The point estimate of instruction duration is -0.003 but again statistically insignificant for science instruction. However, the spillover effect of mathematics is positive and size of 0.032 per minute at a 95% significance level. This means that increasing science instruction duration over the average has no effect but increasing mathematics instruction positively effects the science scores. As TIMSS science questions requires high portion of analytical thinking skills[14], effective mathematics lectures enable students to answer questions with high accuracy. But it does not explain why science lectures' effects are insignificant.

[14] As the questions are based on analyzing a natural process and the "cause and effect" relationship, it is highly conjecturable. For more detail please visit the page of lastest test concept, avaiable online at: https://timssandpirls.bc.edu/timss2019/frameworks/framework-chapters/science-framework/science-content-domains-fourth-grade/

Table 5. 2: Effect of instruction on mean TIMSS science test scores.

	(1)	(2)
Science Instruction	0.001 (0.030)	-0.003 (0.017)
Mathematics Instruction	0.028 (0.024)	0.032** (0.016)
Country Dummies	x	x
All Controls		x
No. of Obs.	19704	11932

Reported numbers are based on weighted linear regressions over one subject (Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

5.1 Results with the First Difference Approach

The research strategy of this study is to form a first difference equation as presented in Equation (4.3) in order to remove unobserved student, teacher, class, and school fixed effects.

In Table 5.3, the results of FD estimation for the general dataset are presented. For the first column, only country dummies are added to the regression. Obtained results are 0.030 for mathematics at 95% significance level, and 0.049 for science at 99% significance level. For the second column, while keeping country dummies, the control variables are also added to the regression. The effects are now increased to 0.050 for mathematics at 99% significance level and 0.094 for science at 99% significance level. Based on the second column, 0.050 points per minute/week means 3 points(60 times 0.050 as the result shifts from minute to hour) per hour/week on average for mathematics; 0.094 points per minute/week means 5.64 points per hour/week for science. These results are similar to what Lavy (2015) and following. Cattaneo et al. (2016) presented. Their estimates were based on the percentage change of 1 SD of the scores for 1 hour increase of instruction.[15] If a comparison with previous regression tables (Table 5.1 and Table 5.2) is held, mathematics instruction effect has decreased

[15] Lavy(2015) found 5.8% SD change in case of an hour increase in instruction by taking the study's whole PISA sample into consideration by using his FE Model setting as a general result for math, science and language subjects. As 1 SD equals to 84.4 for the dataset, results imply 4.89 points increase. Cattaneo et al. (2016) found 5.9% change with similar setting but using study's Swiss students PISA sample. As 1 SD equals to 85.3 for the dataset, results imply 5.03 points increase.

from 0.082 to 0.050 while in science the effect now has positive sign as 0.094 and significance. These changes suggest that there was a large correlation between the error term and the dependent variable because of the unobserved fixed effect and the endogeneity. Application of FD regression reduced the effect of mathematics instruction as it was overestimated in Equation (4.1) and increased the effect of science instruction as it was underestimated with insignificance in Equation (4.2). Also, the results found in the first difference estimation are very similar to what Rivkin (2015) has found. In their paper, one additional minute of instruction's effect was significantly 0.051 as the marginal coefficient for the lectures whose durations are between 240-279 per week by considering 2 subjects included in the study (language arts and math) on average.

On the other hand, it is important to mention that the effectiveness of class hours is the key accelerator of the learning process resulting in higher achievement scores. Diminishing returns are also available in the education production function as students and teachers get tired of instruction. As Aronson et al. (1999) implied quality of teaching including effective use of instructional time is the key input.

Table 5. 3: Effect of instruction on mean TIMSS test scores (FD Method).

	(1)	(2)
Mathematics Instruction	0.030**	0.050***
	(0.014)	(0.012)
Science Instruction	0.049***	0.094***
	(0.016)	(0.016)
Country Dummies	x	x
All Controls		x
No. of Obs.	19704	11907

Reported numbers are based on first difference estimates over two subjects (Math and Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Science instructional time effect has been arranged to make it comparable with math instructional time effect. Details are on the 'Empirical Results' section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

5.2 Country Based Results with the First Difference Approach

In Table 5.4, the results of FD estimation for unique countries in the whole dataset are presented. For the first column, the equation only includes achievement scores and explanatory variables as the previous ones. In the header, country names are placed, and each country has its estimation results below. In comparison to Canada, other countries' estimations are insignificant and contain much higher standard errors. In the case of Canada, similar to general results but stronger effects exist such as mathematics instruction effect is 0.061 and science instruction effect is 0.133 with 99% significance level for Equation (4.2). If a comparison is held between these and those we saw in general results, mathematics instruction's effect is around 20% higher while science's effect is higher more than 40%. Canada shows significantly stronger positive productivity for instruction time while other countries remain inefficient. There are two results of this outcome: First, as the Canadians are almost half of the whole dataset, the general results high likely are influenced by Canada. Second, the difference between Canada and other countries should be investigated. In this paper, an analysis based on interpretation and in a quantitative manner, logit and probit regressions are formed to have an overview of these four countries' educational systems and to illuminate possible reasons for the difference in the effect of instructional time.

Table 5. 4: Effect of instruction on mean TIMSS test scores for all countries (FD Method).

	Canada		Azerbaijan		Slovak Republic		Spain	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Mathematics Instruction	0.028	0.061***	-0.029	0.055	-0.023	-0.004	0.004	0.006
	(0.018)	(0.014)	(0.091)	(0.119)	(0.045)	(0.044)	(0.025)	(0.018)
Science Instruction	0.108***	0.133***	-0.248	-0.340	0.041	0.053	0.001	0.002
	(0.040)	(0.025)	(0.312)	(0.394)	(0.045)	(0.034)	(0.019)	(0.013)
All Controls		x		x		x		x
No. of Obs.	8516	4997	3864	1485	2608	2079	4716	3346

Reported numbers are based on first difference estimates over two subjects (Math and Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Science instructional time effect has been arranged to make it comparable with math instructional time effect. Details are on the 'Empirical Results' section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

5.3 Country Specific Education Systems

As all countries have different histories and different geological presences, because of them and other possible reasons, their education structure is different. In this regard, it is important to distinguish and explain their education system characteristics while interpreting the differences in the effect of instruction duration. Glewwe et al. (2020) presented an education production function that takes school and teacher characteristics, year of schooling, child characteristics, household characteristics, and additional inputs such as school material supply and child's attendance, etc. as inputs, and achievement as output. In Table 5.5 below, school and teacher characteristics and additional inputs are presented. The number of published citable documents in the education area is added for having an overview of the motivation behind the educational progress and the quality of educational development in the related country.

Table 5. 5: Important Educational Characteristics of Countries.

Topic/Country	**Canada**	**Azerbaijan**	**Slovakia**	**Spain**
Educational autonomy in primary and secondary education	District-based autonomy with class level attendance to governance: Province based autonomy and decentralized education system exist. Each province has their own ministry of education with their own rules. In addition, "board/district" based differences highly available on educational resources such as professional development or curriculum determination. Also, the teachers may affect her schools' policies by influencing instructional times and lesson plans.	Full central autonomy: Ministry of Education of Azerbaijan is the central body of education. Also, the Cabinet of Ministers plans and delivers strategies on education. (UNESCO,2010/2011)	Region-based autonomy with considering individual schools: Decentralized national education system exists as eight autonomous, self-governing regions (samosprávny kraj) available in the country. Primary schools, preschools, and school facilities are managed by municipalities, while secondary schools are administered by self-governing regions. The Ministry is the national body which develops unified education policies and strategies. Curriculum is determined by Ministry of Education but tailored for individual needs of schools and regions. (TIMSS '19 Slovakia Report)	Limited autonomy: The Spanish government administer the education system regarding legislation, basic structure, and cooperative initiatives among all autonomous communities. (TIMSS '19 Spain Report) According to the OECD 2018 Report, schools providing primary to secondary education have extremely limited autonomy.

Table 5.5(Cont.): Important Educational Characteristics of Countries.

Government expenditure on education in US$ [16]	82,348.74	1,849.48	4,517.19	58,982.77
Government expenditure in US$/population[17]	2153.13	1823.01	829.26	1246.29
GDP% expenditure for education [18]	5.3	2.7	4.0	4.2
Pupil/teacher ratio, primary education[19]	17	15	16	13
Teacher certification requirements	Being a primary grade teacher in the public schools requires certification. This means all teacher candidates should complete one undergraduate degree and with it, separate degree in education. Besides formal educational needs, teacher candidates should have an internship with a mentor supervisor teacher for a specific duration in a real class environment. Additionally, the membership in a province's teachers' association may be required. (National Occupational Classification, Government of Canada) As it is usually the case in Canada, there may be specific requirements of each province to certify a teacher.	Colleges, universities and education faculties is training center for Azerbaijani primary teacher candidates. The education duration in training institutions is 4 years including field training and profession courses. The candidates also should attend a 14-week internship at public schools. (Uzeirli et al, 2021) After finishing the degree, to a placement in public school, candidates should the central examination and interview process (Müəllimlərin İşə Qəbulu, MIQ). According to the UNESCO 2011 Report, teacher education is not well sufficient and enough qualified. (UNESCO,2010/2011)	Primary and secondary grade teachers need to have five-year teacher training program, which follows a necessary teaching internship. (OECD, 2015) All teachers should have university education except the kindergarten teachers. Teacher training students also have school trips. Besides, technical university graduates may also be teacher after completing related pedagogical and philosophical education. (UNESCO,2010/2011)	Teacher candidates need to finish a training program as minimum a bachelor's degree, and a master's degree for secondary school teachers, then pass a highly competitive examination, and after complete a teaching internship. (OECD, 2018)

[16] UNESCO Institute for Statistics, Canada's expenditure is for the year of 2019; Azerbaijan's expenditure is for the year of 2020; Slovakia's expenditure is for the year of 2019; Spain's expenditure is for the year of 2019. Numbers are as millions.
[17] Population data has taken from World Bank, latest values.
[18] World Bank, latest values have been used.
[19] World Bank, latest values have been used.

Table 5.5(Cont.): Important Educational Characteristics of Countries.

The ratio of covered materials in accordance with TIMSS'19 topics [20]	Math	82%	97%	65%	76%
	Science	56%	58%	77%	67%
Number of published citable documents in the education area[21]		37,537	125	2,491	31,432
Teachers' statutory salaries (2020)[22]		40,035	-	16,757	43,135

Before having an interpretation from above table, it is important to analyze an education production function which is shared below by Glewwe et al. (2020):

$$A = a(S, Q, C, H, I)$$

where A is achievement, S is years of schooling, Q is the set of school and teacher characteristics (in other word, quality), C is pupil characteristics and H is household characteristics that affect learning, and I is educational other inputs that such as pupil attendance to classes and purchases of schoolbooks and other school materials. With another model by Carrol (1963), degree of learning was put at another function. As this paper is investigating the effect of instructional time on achievement, Carrol's model which is based on "time" enables us to distinguish several different "time variables" in the total instructional time. For example, "the allocated time" is inclusive to "engaged in learning time". Carrol's model is presented below:

$$Degree\ of\ Learning: f\left(\frac{(Time\ allocated\ for\ learning) \times (Time\ engaged\ in\ learning)}{(Time\ needed\ to\ learn) \times (Quality\ of\ instruction \times Abiity\ to\ understand)}\right)$$

By looking at these models and the given important educational characteristics of countries in Table 5.5, there may be several explanations for why the instructional time effect is significant and positive only in Canada rather than in other countries. As there is no measure in literature, to our knowledge, for engaged learning time of students, which is also counted as an unobserved characteristic in this study, three points are found as possible reasons for the positive significance of Canada's instructional time on educational output. First, Canada has more knowledge of education in terms of academic research output. This fact reflects the attitude of the academia of the nation

[20] Only primary level subjects were included to ratio. The values are based on teacher survey of TIMSS 2019. Detail information and the status of other countries can be found at http://timss2019.org/download

[21] The time interval for counted documents is 1996-2021.

[22] Used values are from OECD (2021), Annual primary school teachers' salaries; as the section of USD, PPP.

on the education system. In parallel, it can be counted as a quality measure of foundations of education. Increased quality leads more efficient lecture hours resulting higher degree of learning based on the Carrol model. Second, Canada's educational system on primary schools is one of the most decentralized/self-determinative system in this group. A similar result published by Lavy (2015). The study declares that the productivity of instructional time is higher in schools whose extensive autonomy for several aspects of educational process including teacher hiring and firing decisions and budget formation. Most significantly, teachers' voice on instructional hours enables them to arrange the lecture times in a more effective way as they observe and know the characteristics of their classes/students.[23] For example, a teacher may be aware of out-of-focus times of the whole class and arrange the time plan based on this observation. Third, another probable reason is curriculum fulness within allocated time. According to Education at Glance Report of 2019, Canada allocates 5518 hours of instruction, Spain allocates 4750 hours of instruction and Slovak Republic allocates 2768 hours of instruction in primary level as a necessity over the country, annually. Besides, Canada has 9 months interval, Spain has 8.5 months interval, Slovak Republic has 9 months interval and Azerbaijan has 8 months interval consisting of the beginning and ending of school year. According to this time span for countries and the ratio of covered topics, curriculum fulness in the allocated instruction time for a year would be another reason for the productivity of instructional time on achievement. In both subjects, Canada has the lowest coverage of topics with the highest allocated time. Also, Canada's instructional times in the dataset is higher than almost all other countries. The fact can be seen above in Table 3.1. That means Canada allocates more time on one unique topic. These three features imply that higher engaged learning time, higher quality of instruction and self-deterministic structure at primary schools exist in Canada. Overall results with positive effect of additional instructional time in both subjects.

To support or withdraw the interpreted findings above, within the bounds of the possibility of the TIMSS 2019 Dataset, it is possible to examine some of the key metrics by econometric methods. As presented in Table 5.5, only pupil-teacher ratio and curriculum fulness in allocated time can be extracted from the dataset. Based on

[23] Managing instructional time effectively is one of the key elements that increase learning. (Pashler et al. (2007). Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. National Center for Education Research, page 3)

that fact, below, Equation 5.1 and Equation 5.2 present respectively logit and probit models which have the dependent values as the given productivity of addition of instruction time on achievement. The dependent variable is 1 for Canada and 0 for other countries as general results indicate. And the right-hand side values are the curriculum fulness in TIMSS 2019 and pupil teacher ratio.[24]

$$P(Y_i = 1|X_{1sc}, X_{2sc}) = \frac{\exp(\beta_0 + \beta_1 X_{1sc} + \beta_2 X_{2sc})}{1 + \exp(\beta_0 + \beta_1 X_{1sc} + \beta_2 X_{2sc})} \quad (5.1)$$

$$P(Y_i = 1|X_{1sc}, X_{2sc}) = \Phi(\beta_0 + \beta_1 X_{1sc} + \beta_2 X_{2sc}) \quad (5.2)$$

Where Y_i is the response of significant marginal increase or decrease of achievement score to the instructional time additions of the country i, and Y_i is 1 if the country has productivity in terms of additional instructional time, otherwise 0. X_{1sc} is curriculum fulness within limited time of related school of related class in TIMSS 2019 topics, X_{2sc} is pupil teacher ratio, in other words class size, of related school of related class. $\Phi()$ is the cumulative standard normal distribution function. Marginal effects are calculated after the calculation of regression coefficients via logit and probit regression. Table 5.6 and Table 5.7 presents the results for both regressions on the subjects.

Table 5. 6: Logit and Probit Model Results For The Effect of Mathematics Curriculum Fulness on Added Instruction Time Productivity.

		Logit Model	Probit Model
Mathematics Curriculum Fulness	Coefficient	-0.013***	-0.008***
		(0.003)	(0.001)
	Marginal Effect	*-0.0032*	*-0.0032*
Class Size	Coefficient	-0.004	-0.002
		(0.010)	(0.006)
No. of Obs.		1155	1155

Reported numbers are based on logit and probit regressions. Marginal effects are calculated via Logit/Probit model average marginal effects. Marginal effect of "Class Size" is not calculated as it is not statistically significant.
(*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10 and there is no p-value for calculated marginal effects.)

[24] Detail information about variables can be found in "Data" section.

Table 5. 7: Logit and Probit Model Results For The Effect of Science Curriculum Fulness on Added Instruction Time Productivity.

		Logit Model	Probit Model
Science Curriculum Fulness	Coefficient	-0.039***	-0.023***
		(0.003)	(0.001)
	Marginal Effect	*-0.0084*	*-0.0083*
Class Size	Coefficient	-0.016	-0.009
		(0.011)	(0.007)
No. of Obs.		1139	1139

Reported numbers are based on logit and probit regressions. Observations are on class level and same dataset that contains observations from Canada, Azerbaijan, Slovak Republic, and Spain is used. Marginal effects are calculated via Logit/Probit model average marginal effects. Marginal effect of "Class Size" is not calculated as it is not statistically significant. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10 and there is no p-value for calculated marginal effects.)

By running the logit and probit regressions, it is tested that the response of significant marginal increase or decrease of achievement score to the instructional time additions whether or not depends on curriculum fulness and class size of related class c. Upon the significant coefficient of curriculum fulness in both subjects and insignificant coefficient of class size, we can conclude that more instruction in limited time decreases the significance of additional hours on achievement of the pupil. 1% increase in science curriculum fulness with the limited time, causes 0.84% - 0.83% decrease in the chance of being productive of the effect of additional instruction time on success. Figure 5.1 and Figure 5.2 shows the logit and probit curves' goodness of fit respectively.

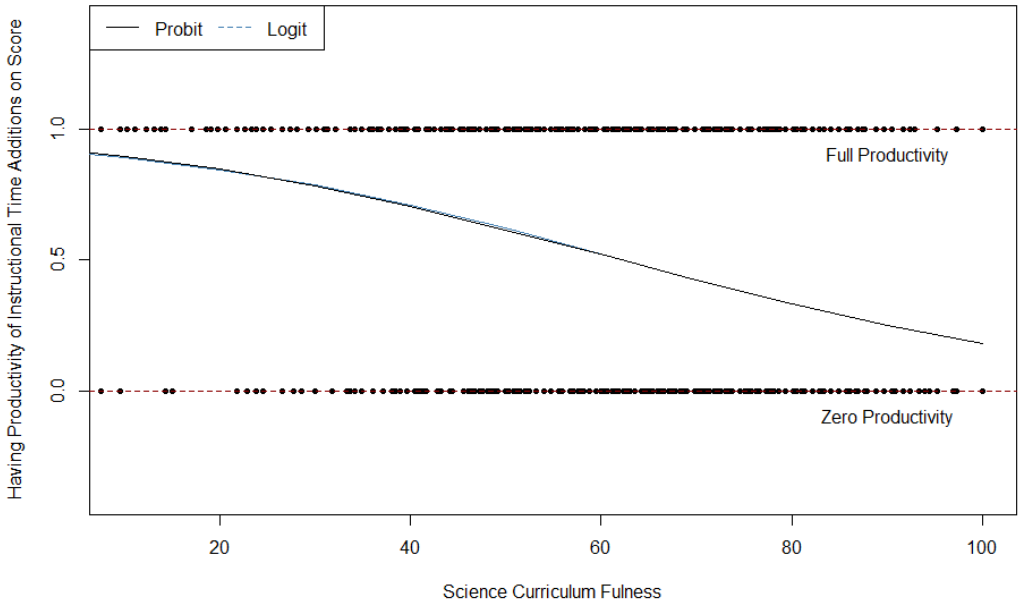

Figure 5. 1: Logit and Probit Model Graph For The Effect of Science Curriculum Fulness on Added Instruction Time Productivity of Science Success

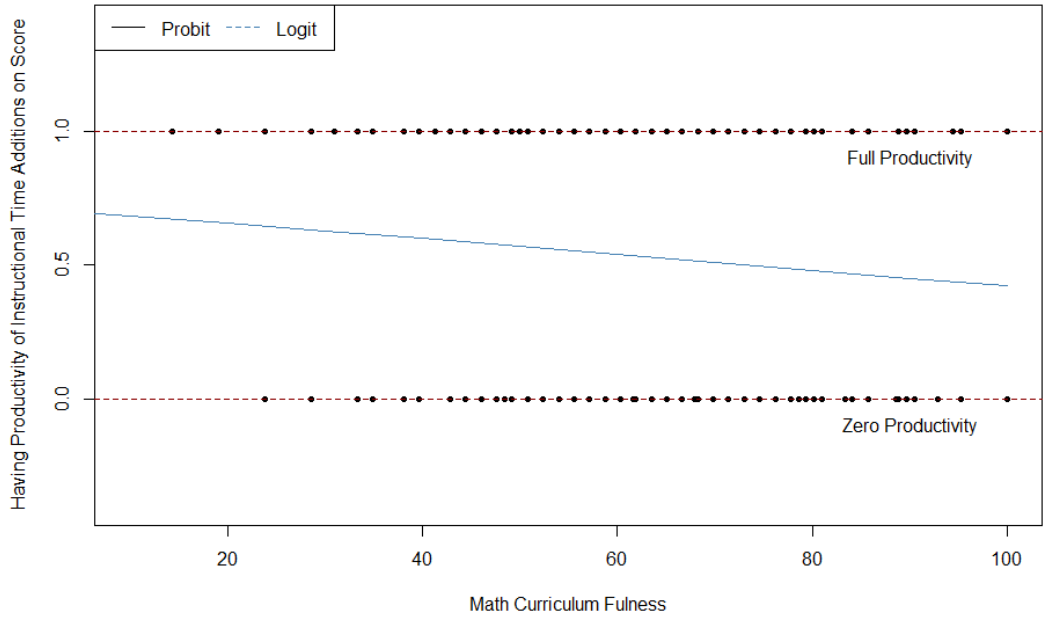

Figure 5. 2: Logit and Probit Model Graph For The Effect of Mathematics Curriculum Fulness on Added Instruction Time Productivity of Success

6. RESULTS ON A SPECIAL CASE: CANADA

The Table 6.1 and Table 6.2 present results respectively the effect of mathematics and science instructions on the subjects for 4th grades from estimated Equation (4.1) and Equation (4.2) by general linear regression for Canada(. The first column in Table 6.1 which does not include any control variables, the point estimate of instruction duration of mathematics lessons is 0.041 and statistically significant at 90% significance level. This means that 4th grade pupils who have 1 hour more in their weekly curriculum get 2.46 points better, on average. Besides, the spillover effect of science instruction is -0.088 with 95% significance level. The second column in Table 6.1, includes all background controls which is presented in the "Data" section. The point estimate of instruction duration is much higher, 0.093 at 99% significance level. This means a stronger effect which indicates that 1 more hour in the weekly curriculum results in 5.58 more points. However, the science spillover effect negatively hits by 0.108 meaning that increasing science instruction duration by 1 hour/week ceteris paribus, results in 6.48 points decrease.

Table 6. 1: The effect of minutes of mathematics instruction on TIMSS scores in Canada.

	(1)	(2)
Mathematics Instruction	0.041*	0.093***
	(0.023)	(0.020)
Science Instruction	-0.088***	-0.108***
	(0.044)	(0.027)
All Controls		x
No. of Obs.	8516	5018

Reported numbers are based on weighted linear regressions over one subject (Math). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

In the first column in Table 6.2, the point estimate of instruction duration of science lessons is 0.021 and statistically insignificant. Besides, the spillover effect of

mathematics instruction is also insignificant. The second column in Table 6.2 includes all background controls presented in the "Data" section. The point estimate of instruction duration is 0.022 but again statistically insignificant for science instruction. However, the spillover effect of mathematics is positive and size of 0.032 per minute at 95% significance level. This means that increasing science instruction duration over the average has no effect but increasing mathematics instruction positively effects the science scores. In the previous part, Section 5.2, at Table 5.4, those that present the first difference estimates after the removal of potential biases are already outlined for Canada. The mathematics instruction effect has decreased from 0.093 to 0.061 due to the potential biases. For the science, the effect is 0.133 with significance rather than the insignificance results. The following section enables us to distinguish if any heterogeneity across the sample exists.

Table 6. 2: The effect of minutes of science instruction on scores in Canada.

	(1)	(2)
Science Instruction	0.021 (0.016)	0.022 (0.022)
Mathematics Instruction	0.013 (0.017)	0.032** (0.016)
All Controls		x
No. of Obs.	8516	5008

Reported numbers are based on weighted linear regressions over one subject (Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

6.1 Heterogeneity Checks: Immigrants, Girls, and Boys

6.1.1 Immigrants

It is important to analyze different subgroups in the country that has productivity(responds positively to instructional time) to the instructional time because educational policies are better to be allocated based on these heterogeneities. For example, schools that include more students with immigrant background than others have different characteristics in the class and school environment. In this regard, the first difference regression results in this section, in Table 6.3, include only the

participants who have immigrant background. Immigrant status determined from "Born in Country" data that collected by TIMSS'19. If the answer is "no", the pupil was counted as an immigrant. 1054 of the 8516 students from Canada are counted as immigrants. In the second column which all the controls are added to Equation (4.3), the effect of mathematics instruction time is 0.053 with 90% significance, also the effect of science instruction time is 0.159 with 99% significance. Canada's overall results presented in Table 5.4, the effect of mathematics instruction time was 0.061 which is higher than that for immigrants. As a result, immigrants show lower productivity to the instructional time increases in mathematics. For science, the effect was 0.133 in Table 5.4, but the effect is 0.159 for the immigrants. In science, lecture duration productivity is positive and higher than the overall sample. As the concept of mathematics, building relation between intangible terms may require more mental attendance by student herself when compared to science which is based on observable experiments and is more extravert. Upon that, immigrants may need more help while covering science topics as it has more verbal and visual concepts than mathematics. Having more instruction time enables them to be engaged more with the concepts and learn more in both subjects but with a higher impact for science lectures.

Table 6. 3: The effects of science and mathematics instructions on the achievement of immigrants in Canada (FD Method).

	(1)	(2)
Mathematics Instruction	0.027 (0.035)	0.053* (0.030)
Science Instruction	0.086 (0.067)	0.159*** (0.030)
All Controls		x
No. of Obs.	1054	685

Reported numbers are based on first difference estimates over two subjects (Math and Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Science instructional time effect has been arranged to make it comparable with math instructional time effect. Details are on the 'Emprical Results' section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

6.1.2 Girls and boys

Table 6.4 presents the results for girls and boys, respectively. Based on the results in the second column, the effect of mathematics instruction time for girls and boys,

respectively, are 0.065 and 0.055. And the effect of science instruction time for girls and boys, respectively, are 0.142 and 0.122. Upon the results, in both sections girls are more positively affected by (more productive) the increase of lecture time. The obtained results for all groups are similar to the findings of Lavy (2015) as there are higher effects for girls and immigrants.

Table 6. 4: The effects of science and mathematics instructions on the achievement of girls and boys, separately in Canada (FD Method).

	Girls		Boys	
	(1)	(2)	(1)	(2)
Mathematics Instruction	0.029	0.065***	0.029	0.055***
	(0.022)	(0.017)	(0.023)	(0.019)
Science Instruction	0.115***	0.142***	0.092**	0.122***
	(0.038)	(0.029)	(0.045)	(0.028)
All Controls		x		x
No. of Obs.	3894	2541	3951	2509

Reported numbers are based on first difference estimates over two subjects (Math and Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Science instructional time effect has been arranged to make it comparable with math instructional time effect. Details are on the 'Empirical Results' section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

6.1.3 Is there an optimal time?

This section tries to explain the optimum instruction time for the sample. One may ask if the lecture time effect is positive and continuous why all schools do not increase the time to the value giving the best score for them? In this paper, to address this question and go into more detail in the investigation, "the possible optimum lecture time" for both subjects is studied across the Canadian sample. To achieve this, instructional durations are grouped concerning the intervals which are separated by 10% increase in the mean. As the effect is positive in both subjects, intervals are made above the "mean" lecture durations. For mathematics subject, the mean was approximately 311 minutes and for the science subject, the mean was approximately 136 minutes. By doing that, one needs to be sure that the intervals are not so small for the sampling variability.

To use these intervals on the regression, dummy variables are employed for each. If the i^{th} student's lecture time on j^{th} topic lays down on specific interval presented

below, the optimum time dummy takes "1", else "0". As there is one more group exists namely, “below the mean”, all optimum time dummy variables are put to the regression. By doing that, left-out dummy is selected as the ones that below the mean and the regression does not contain errors by "dummy trap". Also, Cattaneo et al. (2016) used similar setting by creating categorical variables which are made of specific intervals of weekly instructional lecture hours. The Figure 6.1 shares all the dummies in a chart below.

136 < Optimum Science Instruction 1 <= 149	311 < Optimum Math Instruction 1 <= 341
149 < Optimum Science Instruction 2 <= 162	341 < Optimum Math Instruction 2 <= 371
162 < Optimum Science Instruction 3 <= 175	371 < Optimum Math Instruction 3 <= 401
175 < Optimum Science Instruction 4 <= 188	401 < Optimum Math Instruction 4 <= 431
188 < Optimum Science Instruction 5 <= 201	431 < Optimum Math Instruction 5 <= 461
201 < Optimum Science Instruction 6 <= 214	461 < Optimum Math Instruction 6 <= 491
214 < Optimum Science Instruction 7 <= 227	491 < Optimum Math Instruction 7 <= 521
227 < Optimum Science Instruction 8 <= 240	521 < Optimum Math Instruction 8 <= 551
240 < Optimum Science Instruction 9 <= 253	551 < Optimum Math Instruction 9
253 < Optimum Science Instruction 10 <= 266	
266 < Optimum Science Instruction 11 <= 279	
279 < Optimum Science Instruction 12	

Figure 6. 1: Optimum time dummies (in minutes)

The results in Table 6.5 presents the results of regression including each dummy variable instead of real mathematics instructional time. Only several dummies’ estimates were obtained with significance. However, all except the first dummy have a positive sign. In detail, the most beneficial time when compared to the left out dummy category is the interval of "521-551" which is symbolized as “Optimum Math Instruction 8”. One should ask if there exists a specific school effect. It is controlled that there are several, 3, schools exist in that interval.

The results in Table 6.6 presents the results of regression including each dummy variable instead of real science instructional time. Only several dummies’ estimates are obtained with significance, namely the 9th and 10th dummies. However, all except the first, third, fourth, and eleventh dummies have a positive sign. In detail, the most beneficial time is the interval of "254-266" which symbolized as “Optimum Science Instruction 10”. One again should ask if there exists a specific school effect. It is controlled that there are several, 2, schools for that interval. These investigations are

additions to what is mainly focused through the paper, but optimum lecture time should be investigated more by in-depth analysis. In both regression results, the highest values are not the ones that contribute more to the scores. This can be interpreted as diminishing returns are affirmed.

Table 6. 5: Regressing optimum time dummies on the mathematics achievement score.

	Results
Optimum Mathematics Instruction 1	-3.833 (6.826)
Optimum Mathematics Instruction 2	18.692*** (5.798)
Optimum Mathematics Instruction 3	13.727* (8.109)
Optimum Mathematics Instruction 4	11.142 (7.290)
Optimum Mathematics Instruction 5	19.112*** (7.983)
Optimum Mathematics Instruction 6	8.561 (8.918)
Optimum Mathematics Instruction 7	6.922 (9.887)
Optimum Mathematics Instruction 8	49.488*** (13.483)
Optimum Mathematics Instruction 9	45.401** (19.416)
Science Instruction	0.119*** (0.027)
All Controls	x

Reported numbers are based on general regression estimates over one subjects (Math). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Science instructional time effect has been arranged to make it comparable with math instructional time effect. Details and definitions of each interval (written as 'Optimum Math Instruction') are on the 6. Section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

Table 6. 6: Regressing optimum time dummies on the mathematics achievement score.

	Results
Optimum Science Instruction 1	-4.308 (9.184)
Optimum Science Instruction 2	2.872 (5.829)
Optimum Science Instruction 3	-20.909 (22.364)
Optimum Science Instruction 4	-4.471 (7.103)
Optimum Science Instruction 5	3.772 (6.216)
Optimum Science Instruction 6	17.736 (12.617)
Optimum Science Instruction 7	3.572 (9.510)
Optimum Science Instruction 8	3.313 (9.474)
Optimum Science Instruction 9	32.325** (14.654)
Optimum Science Instruction 10	38.564*** (11.026)
Optimum Science Instruction 11	-5.127 (20.535)
Optimum Science Instruction 12	1.037 (7.122)
Mathematics Instruction	0.032* (0.017)
All Controls	x

Reported numbers are based on first difference estimates over one subject (Science). Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Details and definitions of each interval (written as 'Optimum Science Instruction') are on the 6. Section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

6.1.4 Placebo test

Placebo tests and robustness checks are widely used in economics and particularly education economics. In this paper, to be sure that identification strategy is significant and accountable, a placebo test is performed by employing a randomization technique. The main question is that "If the lecture times are chosen randomly, would we see the same results?". To answer that question, 100 random synthetic instructional duration values are created by a random number generator in the statistical package. The borders of synthetic values are selected as the minimum and maximum lecture durations of the TIMSS'19 database used in this research to have a relevant number set like the real ones. In total 50 regressions were run as half of the values are used instead of real science instructional time values and the other half is used instead of real mathematics instructional time values. Each regression is based on Equation (4.3) which is a setting for the First Different model. As in previous ones, regressions include 1 random lecture time for science instructional time and 1 random lecture time for mathematics instructional time. In total, there were 100 random lecture time variables. From these 50 regressions and 100 random lecture time values only 2 of them gave significant results. (2/100) As this ratio is lower than the 5/100(95% significance level), we are confident that the lecture time influences student scores. While it underlines lecture times' significance, the direction of the relationship is not provided. One of these 50 regression results is presented in Table 6.7 below. Others are available upon request to the author.

Table 6. 7: Placebo effect of random instruction minutes on TIMSS achievement scores in Canada (FD Method**).**

	Results
Mathematics Instruction(r)	0.002 (0.002)
Science Instruction(r)	0.000 (0.002)
All Controls	x

Reported numbers are based on first difference estimates over two subjects (Math and Science) but generated random instructional time variables (by considering minimum and maximum values of real variables and by using uniform random distribution) are relocated for real instructional time variables and signed with (r) mark. Estimates and standard errors are obtained using the five plausible values of test scores as required by TIMSS 2019 User Guide for the International Database 2nd Edition. Standard errors in parentheses are clustered at the class level. Total student weights(totwgt) are calculated and used from TIMSS'19 Database. Random science instructional time effect has been arranged to make it comparable with random math instructional time effect. Details are on the 6. section. (*** as p-values<0.01, ** as p-value<0.05, * as p-value<0.10)

7. CONCLUSION

Additions to instruction time can be seen as a policy that enhances learning. In this study, the instructional time effect was studied over four countries and education systems, and one country in-depth. Important findings can be listed as instructional time did not have the same effect for all education systems and the effect of instructional time was positive in mathematics and science with high variance in an instructional time productive country. However, the size of the effect was bigger in the subject of science. Moreover, heterogeneity-based differences were obtained.

In the first phase of the study, the whole dataset including four different countries and education systems was studied. In overall regression, positive and significant effect of instructional time was found for the whole dataset consisting of over 19,000 observations in both subjects, science and mathematics. Then, with a dubious approach, country-specific research was carried. The significance of the effects differed among different educational systems and countries. The difference was analyzed in-depth by both interpretation and quantitative methods including logit and probit regressions. Correspondingly, the productivity of achievement to instructional time additions over the countries was studied. As a unique approach to our knowledge, logit and probit regressions was formed by employing explanatory variables as curriculum fulness with limited time and teacher-pupil ratio. Moreover, it is determined that adequate time on each topics during class hours is key for instructional time productivity of achievement while the teacher-pupil ratio do not have any effect. Namely, a 1% increase in science curriculum fulness with limited time, causes a 0.84% decrease in the chance of being efficient in the effect of additional instruction time. Similarly, a 1% increase in mathematics curriculum fulness with limited time, causes a 0.32% decrease in the chance of being efficient of the effect of additional instruction time. Also, school autonomy and academic research for quality of education were interpreted as the two measures that differed among countries as possible pillars of instructional time productivity of achievement.

In the second phase, as an instructional time-productive country, Canada was investigated in detail analysis. By controlling a rich information set for pupil, school, classroom, and teacher backgrounds, findings presented a significant positive relationship between achievement scores and instructional time variables using linear regression and first difference linear regression as the standard errors were lower for the latter. An additional minute of mathematics lecture resulted in a 0.061 point marginal score increase on the mathematics test. An additional minute of science lecture resulted in a 0.133 point marginal score increase on the science. Both result are very similar to what Lavy (2015) and Cattaneo et al. (2016) have found.[25]

Thirdly, heterogeneity-based differences was examined. Toward this purpose, three subgroups have been formed and studied: immigrants, girls, and boys. Immigrants showed lower tolerance and a lower significance level in mathematics instructional time increases when compared to science instructional time increases however the effect was positive for both. The second subgroup, girls, showed a significantly higher effect of additional instructional time in both subjects when comparared to the boys and immigrants. In the subgroup of boys, the effect was positive for both subjects but smaller than the other subgroups. The effect of increase in science instruction time was higher than that in mathematics instruction time for all subgroups. Moreover, an optimal duration for lectures was investigated. Possible results for diminishing returns were found which means that more instructional hours do not always result in more success. The pupils who have the highest lecture time were not found at the top of the achievement scale.

Finally, the exact causation behind instructional time and achievement has complexity. Adequate time on each topic during class hours is found foremost for instructional time productivity of achievement for diverse educational systems. In an instruction time-productive education system, heterogenous diversity among pupils based on gender and immigrant status affects the utility of additional instructional time meaning that girls and immigrant benefits more. These results also suggest that in-depth analysis should be done based on the pupil and educational system characteristics of the country before policy implementation.

[25] This comparison is held by considering standard deviation-based percentage increase per hour. Details can be seen at Lavy (2015) and Cattaneo et al. (2016).

REFERENCES

Baker, D.P., Fabrega, R., Galindo, C. et al. (2004). Instructional Time and National Achievement: Cross-National Evidence. Prospects 34, 311–334

Blank, Rolf K. (2013). Science Instructional Time is Declining in Elementary Schools: What Are the Implications for Student Achievement and Closing the Gap?, Science Education 97 (6): 830–847.

Britannica. (2013). T. Editors of Encyclopaedia. "Catholic Monarchs." Encyclopedia Britannica, December 2.

Cattaneo, Maria & Oggenfuss, Chantal & Wolter, Stefan. (2017). The more, the better? The impact of instructional time on student performance. Education Economics. 25. 1-13.

Dobbie, Will, and Roland G. Fryer Jr. (2011**).** Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." American Economic Journal: Applied Economics, 3 (3): 158-87.

Eric Eide and Mark Showalter. (1998). The effect of school quality on student performance: A quantile regression approach, Economics Letters, 58, (3), 345-350

Farbman, David A., David J. Goldberg, and Tiffany D. Miller. (2014). Redesigning and Expanding School Time to Support Common Core Implementation. Washington, DC: Center for American Progress. National Center on Time & Learning.

Fryer Roland G. Jr. (2011). Injecting Successful Charter School Strategies into Traditional Public Schools: A Field Experiment in Houston.

Gromada, A. and C. Shewbridge. (2016). Student Learning Time: A Literature Review, OECD Education Working Papers, No. 127, OECD Publishing, Paris.

Jensen, Vibeke M. (2013). Working Longer Makes Students Stronger? The Effects of Ninth GradeClassroom Hours on Ninth Grade Student Performance. Educational Research 55 (2): 180–194.

Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries. The Economic Journal 125 (588): F397–F424.

Long, D. A. (2014). Cross-National Educational Inequalities and Opportunities to Learn: Conflicting Views of Instructional Time. Educational Policy, 28(3), 351–392.

Lee, J. and Barro, R.J. (2001). Schooling Quality in a Cross–Section of Countries. Economica, 68: 465-488.

Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). Methods and Procedures: TIMSS 2019 Technical Report. Retrieved from Boston College, TIMSS & PIRLS International Study Center

OECD (2011). Education at a Glance 2011: OECD Indicators, OECD Publishing, Paris.

OECD (2019), Education at a Glance 2019: OECD Indicators, OECD Publishing, Paris.

Pischke, J.-S. (2007). The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years. The Economic Journal 117 (4): 1216–1242.

Pashler et al. (2007). Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. National Center for Education Research

Rivkin, S.G. and Schiman, J.C. (2015), Instruction Time, Classroom Quality, and Academic Achievement. Econ J, 125: F425-F448.

Rønning, M. (2010), Homework and pupil achievement in Norway Evidence from TIMSS. Statistisk sentralbyrå /Statistics Norway, Reports 1/2010.

Shaw, Melanie & Chametzky, Barry & Burrus, Scott & Walters, Kelley. (2013). An Evaluation of Student Outcomes by Course Duration in Online Higher Education. Online Journal of Distance Learning Administration. 16. 1-22.

Satoh H, Tamalu F, Hirosawa N, Hirasawa H, Nagane M, Saito R, Watanabe SI, Miwa N.(2022)Facilitatory Effect of Extending the Course Duration on Dissemination of Educational Content. Med Sci Educ, May 7;32(3):641-648.

Tiedt Jane A., Owens Joan M., Boysen Susan. (2021)The effects of online course duration on graduate nurse educator student engagement in the community of inquiry, Nurse Education in Practice, Volume 55 ,103164,ISSN 1471-5953,

Walberg, H. J. (1991). Does homework help? The School Community Journal, 1(1), 13–15.

Woessmann, L. (2003). Schooling Resources, Educational Institutions and Student Performance: The International Evidence. Oxford Bulletin of Economics and Statistics 65: 117–170.

Yeşil Dağlı, Ümmühan. (2019) Effect of increased instructional time on student achievement." Educational Review 71 501 - 517.

Url-2 Why Time Matters. (2015). National Center on time and learning, n.d. web. http://www.timeandlearning.org/?q=why-time-matters-0 ,date retrieved 11.03.2023.

Url-1 Globe Newspaper Company, http://archive.boston.com/lifestyle/family/articles/2009/09/28/obama_wants_nations_children_to_spend_more_time_in_school/, date retrieved 11.03.2023.

APPENDIX

1) $\alpha_{mat} > 0$ implies "the effect of mathematics time on mathematics score minus the effect of mathematics time on science score" on mathematics minus science scores. By doing that, the impact of pure mathematics instruction on the mathematics score calculated. The effect of science instruction on mathematics is eliminated.

2) $\alpha_{sci} < 0$ implies "the effect of science instruction time on mathematics score minus the effect of science instruction time on science score" on mathematics minus science scores. The mathematics instruction's effect on science scores is eliminated. The effect of pure science instruction time on the science score is calculated. Moreover, as α_{sci} is expected to be negative if the effect is positive, α_{sci} is multiplied by "-1" in order to make it comparable and interpretable with mathematics score. The size remains same.

3) Some of the instruction time variables which was a weekly measure of instruction time on the studied sample was illogical. The authors corrected some weekly instruction time variables which are between 1 and 10 by multiplying 60. The reason behind it is that the teachers might have thought that the scale is "hour" not "minutes".

CURRICULUM VITAE

Name Surname : Eylül Berke Özgen

EDUCATION:

- **B.Sc. :**2020**,** Yildiz Technical University, Electrics and Electronics Faculty, Electrical Engineering
- **M.Sc.** :2023, Istanbul Technical University, Economics

PROFESSIONAL EXPERIENCE AND REWARDS:

- January 2021-October 2021: Tendering Engineer at Kesir Engineering Company
- November 2021- March 2022: Marketing Responsible at Adlema
- March 2022-Present: Electrical Engineer at Bogazici University

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Ozgen, E.,** Aygun H. A. 2022. The Effect of Lecture Time on Student Outcome, Cross-Country Study with TIMSS Data. *International Graduate Research Symposium- IGRS'22*, June 1-3, 2002, Istanbul, Turkey.