Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Investigation of cross-country differences in student performance in standardized tests: the role of modern and traditional teaching methods

Author: **Bc. Marie Ptáčníková**

Study program: **Economics and Finance**

Supervisor: **Mgr. Barbara Pertold-Gebická M.A., Ph.D.**

Academic Year: **2020/2021**

Declaration of Authorship

The author hereby declares that she compiled this thesis independently; using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 4, 2021 Marie Ptáčníková

Acknowledgments

I would like to express my sincere gratitude to my supervisor Mgr. Barbara Pertold-Gebická M.A., Ph.D., for her expert advice, valuable comments and guidance during the entire process of writing my thesis.

Also, I am thankful to my family and boyfriend for their support and continuous encouragement during my studies. Thank you.

Abstract

There is an ongoing debate about what teaching practices are the most effective ones in order to improve student performance. However, little is known about the impact across countries and literature is highly inconclusive. In this work, we extend the portfolio of countries and provide evidence about the role of modern and traditional teaching practices on students' test scores in 43 countries. Our analysis is performed in two steps and is a typical example of hierarchical linear modelling (HLM). In the first step, we perform student fixed effect method to account for majority of selection issues. We identify a positive, negative or no effect of modern or traditional teaching methods on student performance. These results are priceless for policy makers suggesting that there is no one-fits-all-approach towards modern or traditional teaching methods to order to improve students' test scores. As a great variation is observed, we continue further and investigate what country characteristics could explain these differences across countries. Bayesian model averaging (BMA) method supports us in a model uncertainty and a particular variable selection. Our findings indicate that cultural dimension uncertainty avoidance, which describes country's rigidity in behaviour and institutions, assists in explaining these differences. In particular, our results suggest that uncertainty avoidance is positively related to the estimated effect of modern teaching practices on students' test scores. On the other hand, our findings reveal that uncertainty avoidance is negatively related to the estimated effect of traditional teaching practices on student's performance.

JEL Classification	I2, I21, I25, F68, Z10
Keywords	modern teaching practices, traditional teaching practices, TIMSS standardized tests, international comparison, within-student across-subject approach, student achievement

Abstrakt

Jaké metody výuky ve třídě jsou nejúčinnější, aby se zlepšily výsledky žáků, je velmi diskutované téma. O dopadu napříč zeměmi je však známo málo a literatura poskytuje nejasné závěry. V této práci rozšiřujeme portfolio zemí a analyzujeme roli moderních a tradičních výukových metod na výsledky žáků ve standardizovaných testech ve 43 zemích. Naše analýza se provádí ve dvou krocích a je typickým příkladem hierarchického lineárního modelování (HLM). V prvním kroku provedeme metodu fixního efektu žáka, abychom zohlednili většinu problémů s nenáhodným rozdělením žáků do škol a tříd. Identifikujeme pozitivní, negativní nebo žádný vliv moderních či tradičních výukových metod na výsledky žáků v daných zemích. Tato zjištění jsou neocenitelné pro tvůrce politik, jelikož naznačují, že neexistuje jeden univerzální přístup k moderním nebo tradičním metodám výuky za účelem zlepšení výsledků žáků. Jelikož je pozorována velká variace mezi zeměmi, pokračujeme dále a zkoumáme, jaké charakteristiky zemí by mohly vysvětlit mezinárodní rozdíly. Bayesiánská metoda průměrování modelů (BMA) asistuje s modelovou nejistotou a výběrem konkrétních proměnných. Naše zjištění naznačují, že kultura v dané zemi, přesněji řečeno kulturní dimenze vyhýbání se nejistotě a vztah k budoucnosti, která popisuje rigiditu země v chování a institucích, pomáhá vysvětlit tyto rozdíly. Naše výsledky zejména říkají, že vyhýbání se nejistotám pozitivně souvisí s odhadovaným vlivem moderních metod výuky na výsledky žáků. Naopak, vyhýbání se nejistotám negativně souvisí s odhadovaný vlivem tradičních metod výuky na výsledky žáků.

Klasifikace	I2, I21, I25, F68, Z10
Klíčová slova	moderní metody výuky, tradiční metody výuky, standardizované testy TIMSS, mezinárodní srovnání, metoda within-student across-subject, výsledky žáků

Contents

List of Tables

List of Figures

Acronyms

BMA	Bayesian Model Averaging
BMS	Bayesian Model Averaging software (R-package)
CBEEE	Curriculum-Based External Exit Exam by Bishop (1997)
Cond. Post. Sign	Conditional Posterior Sign
GDP	Gross Domestic Product
HLM	Hierarchical Linear Modelling
IMF	International Monetary Fund
IT	Information Technology
ISCED	International Standard Classification of Education
MCMC	Markov Chain Monte Carlo
OECD	Organization for Economic Co-operation and Development
PIP	Posterior Inclusion Probability
PMP	Posterior Model Probability
PPP	Purchasing Power Parity
SD	Standard Deviation
SE	Standard Error
TIMSS	Trends in International Mathematics and Science Study
UNA	Uncertainty avoidance index
WLS	Weighted Least-Squares

Master's Thesis Proposal

Author:	Bc. Marie Ptáčníková
Supervisor:	Mgr. Barbara Pertold-Gebická M.A., Ph.D.
Defense Planned:	February 2021

Proposed Topic:
Investigation of cross-country differences in student performance in standardized tests

Motivation:
Children all around the world spend many years in schools. Education can promote or hinder development of various skills. As teachers play a key role in educating pupils, current research has focused on investigation of teaching practices used in the classroom. A recent study by Bienteck (2014) examines modern vs. traditional teaching methods in the US. Similarly, Korbel & Paulus (2018) look at impact of modern teaching methods vs. traditional ones on student test results in the Czech Republic, and more specifically, on socio-emotional skills that are very important in later stages of our life. They found that modern teaching practices have positive effect on socio-emotional skills while having no adverse effect on test scores, but that standard practices have no impact on socio-emotional skills or test scores in the Czech Republic. That can be translated that changing the methods of teaching towards more modern methods can have positive impact on pupils' socio-emotional skills without affecting the test scores.

However, little is known about the impact across countries and the evidence is highly inconclusive. For instance, Klaveren (2011) finds that different teaching methods do not impact student performance in the Netherlands. Moreover, Lavy (2016) suggests that the teaching method impacts differently students from different backgrounds in Israel. Inspired by Klaveren (2011) and Lavy (2016), we can assume significant differences across countries.

This thesis aims to contribute to the existing body of literature and rising research on student performance in standardized tests and to identify and potentially explain cross-country differences in the extent to which modern teaching practices are used and their effect on students. Standardized tests such as TIMSS 2007 provide public information about student performance with the aim to improve learning (Martin, Mullis & Foy 2008). The nature of data allows us to explore cross-country variation in educational policies. We will try to identify factors predicting differences in modern vs. traditional teaching practices across countries and their impact on socio-emotional skills and test results of secondary school students across countries. Study by Woessmann (2016) suggests that these differences may be explained by cultural factors, society setting, education system or institutional structures.

Hypotheses:
Based on the motivation part, the hypotheses are as follows. Hypothesis 1 is that the causal effect of modern teaching methods on students' performance differs across countries. And Hypothesis 2 that the causal effect of modern teaching methods on student's socio-emotional skills differs across countries. We will test these hypotheses using the TIMSS 2007 data.

If we find significant cross-country differences in the relationship between teaching methods and students' outcomes, we will investigate them further. Hypothesis 3 is why modern teaching methods have impact on test scores in some countries while not in other ones. And similarly, Hypothesis 4 for socio-emotional skills.

In the regression, the explanatory variables of our interest will be country characteristics. These may include cultural factors, societal setting, educational system or institutional structures (Woessmann 2016). Also, some societies may valuate achievement differently or have different educational systems with (or without) exit exams, early tracking, mandatory nursery school attendance, with a different starting age, a different proportion of students attending private schools, motivate teachers in a dissimilar way, etc. (Hanushek & Woessmann 2011).

1. Hypothesis 1: The causal effect of modern teaching methods on student's performance differ across countries.
2. Hypothesis 2: The causal effect of modern teaching methods on student's socio-emotional skills differ across countries.
3. Hypothesis 3: Why do modern teaching methods have impact on test scores in some countries and not in another countries?
4. Hypothesis 4: Why modern teaching methods promote socio-emotional skills in some countries and not in another countries?

Methodology:

In this thesis, our analysis will take place in two stages. First, we will use the student-fixed effect model (i.e. within-student between-subject strategy) for each country, in line with other studies of this type, including Bietenbeck (2014) and Korbel & Paulus (2018). The outcome variable (test score and self-confidence) of a student in specific subject taught by specific teacher will be regressed on standard vs. modern teaching method, teachers' characteristics, and student fixed-effects. The equation will take the following form.

$$TestScore_{ijtc} = \beta_{0c} + \beta_{1c}traditional_{ijtc} + \beta_{2c}modern_{ijtc} + Z_{jtc}\beta_{3c} + a_{ic} + u_{ijtc},$$

where index *i* stands for a student, *j* for a subject, *t* for a teacher and index c for a given country c. The outcome variable $TestScore_{ijtc}$ of a student in a specific subject taught by a specific teacher is regressed on traditional $traditional_{ijtc}$ vs. modern $modern_{ijtc}$ teaching methods, teachers and class characteristics Z_{jtc}, and student fixed-effects a_{ic} which accounts for any student characteristics determining the outcome variable. The u_{ijtc} is the error term.

We will estimate betas for each country with an index c which then will enter in the second stage as our dependent variable. In the second stage, we will take a full list of $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$ estimates and regress these on country characteristics using weights proportional to the precision with which betas are estimated. The goal will be to identify factors explaining differences in student performance in standardized tests among the selected countries, with the use of weighted least-squares (WLS) method.

Data from the Trends in International Mathematics and Science Study (TIMSS) in 2007 that was collected for fourth and eighth graders in 59 countries will be used. We use the 2007 TIMSS wave because it is the last one including students' reports on teaching practices. The data is publicly available on the survey's website. The data provides information on students' test scores, other students' and teachers' characteristics along with teaching methods collected through the self-reported questionnaires.

Moreover, the students' questionnaires contain data on teaching practices based on which we will construct aggregate class indexes on modern vs. traditional teaching practices in line with Bietenbeck (2014) and Korbel & Paulus (2018). Students were asked to answer on a scale whether a particular teaching method was used in the classroom, separately for math and science. Modern methods are defined as working in groups, having discussions, and using what they learn in daily lives unlike traditional methods which include rote learning and memorizing.

We will work with econometric models including the Fixed Effect model and the weighted least-squares (WLS) estimation method using the R software.

Expected Contribution:

This study aims to contribute to the existing literature and bring new findings about the factors that predict differences in test results and socio-emotional skills among lower secondary school students across countries. We will especially focus on the role of modern vs. traditional teaching practices. The main contribution of this thesis is international comparison (i.e. across-countries). To my best knowledge, there was no similar study investigating cross-country differences.

Differences in educational systems and attitudes towards education are substantial across countries to such an extent that it seems impossible to set and apply one-size-fits-all teaching method all around the world. Therefore, we investigate differences in individual teaching methods and how these methods resonate with educational systems and social standards.

Thus, this study would like to help countries to understand to what extent they should apply modern teaching methods given specific (educational policies) characteristics. Results from this work can be useful to policymakers who set educational policies that encourage socio-emotional skills on international levels, as well as to better understand the role of modern teaching methods in educational systems.

Outline:

The structure of my thesis will be as follows.

1. Introduction: I will introduce the topic and provide motivation why this topic is relevant to be studied as well as I will mention potential contribution of my work to current research
2. Literature review: In this part, a detailed overview of the literature and existing body of research will be provided, including the most relevant studies critically analysing them including their methods used, results found and limitations
3. Data: I will describe my dataset in detail and explain how the data was collected. Furthermore, I will briefly describe how I transform the data into the format needed for the model (e.g. construct modern vs. traditional teaching methods indexes)
4. Empirical analysis: In this part, my model(s) and estimation methods that will be used are introduced. Also, I will explain in detail the fixed effect estimation, its advantages and relevance.
5. Results: I will present and discuss my regressions results and robustness checks.
6. Conclusion: In this section, my findings will be summarized, including possible future research avenues and implications of my results for policymakers and educational sector

Core Bibliography:

1. Bietenbeck, J. (2014): “Teaching Practices and Cognitive Skills.” *Labour Economics* **30:** pp. 143-153. doi:10.1016/j.labeco.2014.03.002
2. Hanushek, E. A. & L. Woessmann (2011): “The economics of international differences in educational achievement.” *Handbook of the Economics of Education* **3**: pp. 89-200
3. IEA (2009): “TIMSS.” Available at: https://www.iea.nl/index.php/data-tools/repository/timss
4. Klaveren, C. V. (2011): “Lecturing Style Teaching and Student Performance.” Economics of Education Review **30 (4):** pp. 729–739. doi:10.1016/j.econedurev.2010.08.007
5. Korbel, V. & M. Paulus (2018): “Do teaching practices impact socio-emotional skills?“ *Education Economics,* **26(4):** pp. 337-355. https://doi.org/10.1080/09645292.2018.1460320.
6. Lavy, V. (2016): “What Makes an Effective Teacher? Quasi-experimental Evidence.” *CESifo Economic Studies* **62(1):** pp. 88-125. https://doi.org/10.1093/cesifo/ifv001
7. Martin M.O., I. V. Mullis & P. Foy (2008). “TIMSS 2007 International mathematics report. Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.” *TIMSS 2007*, Chestnut Hill, TIMSS & PIRLS international Study Center, Boston College. Available at: https://timssandpirls.bc.edu/TIMSS2007/context.html
8. Woessmann, L. (2016): “The importance of school systems: Evidence from international differences in student achievement.“ *Journal of Economic Perspectives* **30(3):** pp. 3-32.

1 Introduction

There is an ongoing discussion among parents, teachers and researchers about what teaching practices are best and most effective ones to be used in classrooms. Nowadays, one can observe a promotion of *modern* teaching methods, such as a more individual approach or working in small groups due to a belief that this specific teaching method can improve student performance (National Research Council 1996; Klaveren 2011; Lavy 2016). However, other countries have opposite view and promote *traditional* teaching methods, such as lecturing or memorization (Walker 2012).

Little is known about the impact of a particular teaching method on student performance and evidence is highly inconclusive. Research differs in various ways, but it generally investigates the impact of traditional teaching practices on student test scores, in comparison to the role of modern teaching practices. Some studies report no effect of both modern and traditional teaching methods on student performance (Klaveren 2011; Korbel & Paulus 2018). Other papers suggest positive effect of traditional teaching methods on student performance (Schwerdt & Wuppermann 2011; Bietenbeck 2014). Lastly, research also identified positive effect of modern teaching methods on student performance (Aslam & Kingdon 2011; Lavy 2016). Thus, one can expect significant differences in the role of modern and traditional teaching methods on student results across countries, which is our first hypothesis.

In this paper, we examine the impact of modern and traditional teaching methods on student performance internationally for 43 countries for which data is available in order to draw more consistent evidence and recommendations for policy makers in order to set up educational policies accordingly. For the country-level analyses, we use data from 2007 TIMSS wave which is the last wave in which students reported the frequency of a particular teaching practice they encountered in a classroom. Our analysis focuses on eight-grade students, because data on teaching practices are more detailed and because fourth-grade students are usually taught by the same teacher which would not allow us to employ within-student across-subject approach. The TIMSS data contains information on test scores, student reported answers and teacher and class characteristics.

Following previous literature, class aggregated indexes for modern and traditional teaching methods are created from the TIMSS student questionnaires (Bietenbeck 2014; Korbel & Paulus 2018). The index can be interpreted as an effective share of a class in which students are taught by using modern or traditional teaching methods in a given subject. We use the within-student across-subject variation as an empirical approach which controls for most of selection issues and student sorting. Furthermore, our model control for rich teacher and class characteristics to further limit any possible unobserved characteristics issues.

Our results provide international evidence which suggest that there is no one-fits-to-all approach towards the use of modern and traditional teaching methods to improve student performance across countries. First, in some countries (10), we find positive effect of modern teaching practices on student performance. Second, in one country we identify a negative effect of modern teaching methods on students' test scores. On the other hand, our results suggest a positive effect of *traditional* teaching methods on student performance in several countries (12). In contrast, we find a negative effect of traditional teaching methods on student performance in one country. Furthermore, in two countries we identify a positive effect of both modern and traditional teaching methods on student performance, but the effect is stronger for traditional teaching methods. Lastly, in almost half of the countries (19) in our sample, we find neither modern nor traditional teaching practices have any effect on student performance.

To discuss our results in a more consistent way, we build on our first-stage country-level analyses and identify factors which could potentially explain the observed across country differences in the impact of modern and traditional teaching methods on student performance. This is our second hypothesis: why these differences across countries exist. Our analysis is performed in two stages and thus, it is a typical example of hierarchical linear modelling (HLM). Estimated coefficients from our first-stage country-level analyses enter our second-stage model as a dependent variable, weighted by standard errors, i.e. we use weighted least-squares method (WLS).

Although there is no previous literature investigating why the differences in the effect of modern or traditional teaching methods on student performance exist, there is literature examining differences in student performance in standardized tests across countries in general. In line with our intuition and existing studies which suggests that the differences can be explained by educational system, institutional structures or cultural factors (Hanushek & Woessmann 2010; Woessmann 2016), we collect unique dataset including various country, educational and culture characteristics. Due to model uncertainty, Bayesian model averaging (BMA) assists us with variable selection and supports our preferred model choice.

Our key finding is that cultural dimension uncertainty avoidance assists in explaining the international differences in the effect of modern and traditional teaching practices on student performance in standardized tests. Our results indicate that uncertainty avoidance positively relates to the effect of modern teaching practices on student performance. We can interpret this that in countries with higher uncertainty avoidance index, i.e. countries which prefer planning ahead and rigid codes of behaviour, should rather promote modern teaching methods in order to improve student performance.

Moreover, looking at the differences in the effect of *traditional* teaching methods on student performance across countries, our findings indicate that uncertainty avoidance is negatively related to the effect of traditional teaching methods on students' test scores. Our results would suggest that policy makers in countries with low uncertainty index will more likely benefit from the promotion of traditional teaching methods in order to improve student achievement in standardized tests. This policy recommendation applies to more open-minded countries which do not require to follow strict behavioural patterns.

This study contributes to existing literature in various ways. First, this is the first study which examined the effect of modern and traditional teaching methods on student's test scores for a wider portfolio of individual countries. As the same methodology, data and teaching methods measurement is used, it allows for an immediate comparison of results between countries and provides a more consistent picture for policy makers. Second, this is a first study investigating why these differences in the effect of modern and traditional teaching practices on student performance across countries exists. Besides, this work contributes to a gap in literature calling for examination of international differences beyond existing concepts (Hanushek & Woessmann 2010), including cultural characteristics. Next, our study assists in better understanding of the role of modern teaching methods in educational system in different countries and provides evidence for policy makers whether they should rather promote modern or traditional teaching methods in order to improve student test scores, based on country characteristics, cultural aspect and educational setting in a given country.

This thesis will be structured as follows. First, we review current literature. In the following chapter, we introduce empirical strategy which is used in this thesis. Then, our data is being discussed. Next, we present and discuss our results. Lastly, conclusion of our findings follows.

2 Literature Review

2.1 Research on modern and traditional teaching methods

Existing literature investigating the effect of modern and traditional teaching practices on student performance is scarce (Klaveren 2011; Bietenbeck 2014; Lavy 2016; Korbel & Paulus 2018). In this study, we extend the portfolio of countries for which the evidence of the effect of modern and traditional teaching practices on student performance exists. Besides, we focus on international comparison and identifying patterns in the existing differences across countries which makes our study unique and represents its main contribution. Furthermore, this study contributes to previous research and suggests that the observed differences in few previous studies are not solely due to the different use of methodology, data or approach towards the measurement of teaching methods. We suggest that the differences across countries may be due to country characteristics: educational systems institutions or cultural factors among others.

Nowadays, one could notice an increased promotion of modern teaching practices towards traditional practices in many countries (Bietenbeck 2014; Lavy 2016). This work provides insights into the ongoing discussion on a relative promotion of modern versus traditional teaching methods in educational systems. Traditional or so-called standard methods might be viewed as teacher-centred methods which focus on memorization, lecturing and drill in acquiring knowledge (Lavy 2016). Likewise, traditional teaching methods may be characterized by instil knowledge and comprehension (Lavy 2016). Opponents argue that lecture-style teaching presents many disadvantages, such as a lack of feedback about learning, inflexibility in a specific student pace, a drop in student attention over the course of a class, that a passive learning solely by listening may be unpreferred method for some students and lastly, a quick forgetting as there is no link to an actual practice (Schwerdt & Wuppermann 2011).

In contrast, modern teaching methods are more student-centred and flexible approaches in the classroom, such as collaboration in groups (Bietenbeck 2014). The methods focus on creative thinking, teaching students how to actually think and are believed to develop further skills, e.g. analytical or critical skills (Resnick 1987; Lavy 2016). Opponents of this teaching style argue that promotion of modern methods can cause that students would not acquire basic knowledge, facts and approaches (Hamilton et al. 2003).

The tendency towards modern teaching methods is observed in various countries across the world. In the US, for instance, the National Standards promote modern teaching methods with focus on group activities and cooperation (National Research Council 1996). In Israel, there is a similar trend. A recent reform recommends a reduction in the use of traditional teaching methods at the post-primary level in 2008 (Lavy 2016). However, the views are far from conclusive on both country and international level. For instance, the English Secretary of Education supports traditional teaching practices and calls for their promotion in schools in the United Kingdom (Walker 2012).

The use of a specific teaching method may be favoured due to a belief that a one particular method out of these has a positive impact on student performance or other skills. However, previous literature investigating the impact of different teaching methods on student performance is inconclusive and studies differ in various ways (Bietenbeck 2014; Lavy 2016; Korbel & Paulus 2018).

In the US, studies suggest a positive impact of traditional teaching methods on student performance while no effect of modern teaching practices. First, Wenglinsky (2002) investigated various practical teaching activities and their impact on test scores. He concludes that teachers play an important role in student achievement and that teaching practices, such as real-world problems or emphasis on practical thinking skills, positively relates to student achievement. Later, Schwerdt & Wuppermann (2011) examined two individual teaching practices and their impact on student performance, namely lecturing (which can be viewed as a traditional method) and problem solving (which can be viewed as a modern teaching method). This was an advancement from previous studies which looked at a general teaching style and its effect on student performance. Schwerdt & Wuppermann (2011) conclude that an increase in the use of the traditional method (lecturing) at an expense of the use of the modern method (problem solving) improves student performance.

The following research extended the definition of modern and traditional teaching methods and identified similar results. Bietenbeck (2014) investigated the next TIMSS wave and found a positive impact of traditional teaching methods on test scores while no effect of modern teaching methods in the US, despite that fact that these two aforementioned studies differentiate in an approach towards complementarities of modern and traditional teaching methods. Furthermore, his findings suggest that traditional teaching methods have a positive effect on knowing, applying and reasoning skills as well.

In addition, Lavy (2016) investigated data from Israel and reported a positive impact of both traditional and modern teaching methods on students' test scores. His results suggested the effect of traditional teaching methods to be larger than the effect of modern methods. More specifically, the effect of traditional teaching methods is found to be specific to students from low socioeconomic backgrounds in Israel. Moreover, Aslam & Kingdon (2011) investigated math and language students in Pakistan. They conclude that students whose teachers tend to challenge them and ask a lot of questions in classes have higher achievement in tests. In wider perspective, one could classify this approach as a modern teaching method. As in the vast majority of the above-mentioned studies, authors also employ the within-student across subject variation, which is the approach we follow as well.

In contrast, some studies found no effect of either traditional or modern teaching methods. Klaveren (2011) investigated students in the Netherlands and found neither lecture-style nor individual-style teaching to have any impact on test scores. In addition, Korbel & Paulus (2018) identified no impact of either modern or traditional teaching methods on student performance in the Czech Republic. The latter study defined teaching methods in the same way as Bietenbeck (2014).

Existing literature tends to access two type of data about teaching practices in a classroom, either from teacher or student perspective (Kunter & Baumert 2006). First, the information on teaching practices is reported by teachers themselves stating how much time they devoted to a particular teaching method in a class. For instance, this approach was followed by Klaveren (2011) and Schwerdt & Wuppermann (2011) who investigated impact of teaching methods on student test scores in the Netherlands and in the US, respectively.

The second stream of literature extracts the information on teaching practices from student reported answers (Bietenbeck 2014; Lavy 2016; Korbel & Paulus 2018). One could prefer this approach, as it eliminates the gap between teacher perception of using a certain teaching method and the actual teaching practices used in a given class (Montgomery & Baker 2007). Furthermore, there is a development in the specificity of available data. In a study by Lavy (2016), author uses data from a national survey where students reported the frequency in the use of particular teaching practices across all subjects in Israel. Later research exploited TIMSS international surveys with more detailed student questionnaires where students report how often a certain teaching activity is used in a classroom separately for each subject (Bietenbeck 2014; Korbel & Paulus 2018).

Besides, there is an ongoing debate on whether modern and traditional teaching practices are in fact substitutes or rather complements. Some studies argue them to be complementarities which assures teacher productivity to remain constant independently of increasing one practice during teaching (Bietenbeck 2014; Korbel & Paulus 2018). This is also the approach we hold in this work.

In contrast, the other stream of literature sees modern and traditional teaching practices rather as substitutes (Schwerdt & Wuppermann 2011; Klaveren 2011). It means that using one particular teaching method more would be necessarily at the expense of the other method, i.e. individual approach to a lecture-style teaching. Thus, these studies include correlation coefficient and perform various robustness models to assure their view is in line with their data.

In addition, Bietenbeck (2014) introduced a balanced teaching index which means that both modern or traditional teaching methods index for a given class was above 25th percentile further supporting his complementary view over traditional and modern teaching methods.

2.1.1 Evolution of teaching practices over time

As we argued above, nowadays, one can notice a tendency towards promotion of modern teaching methods. There is a rising research investigating whether the use of modern teaching practices tend to evolve over time. Smith et al. (2002) investigated trends in teaching methods in math and science classrooms in the US in years 1993 and 2000. They looked at one modern and one traditional teaching method: working in small groups and lecturing, respectively. Teachers were asked how often they applied a particular method in their classes. Authors found no significant change in the use of modern teaching method working in small groups in both math and science. In science, there was found a modest decrease in frequency of traditional teaching method lecturing. However, this is possibly due to the short time period.

Likewise, Bietenbeck (2014) attempted to track the use of modern teaching practice working in small groups. Working in small groups is the only teaching practice which was included across all waves of TIMSS testing preceding author's study. Author found a small positive increase within these years. More specifically, students reported working in groups in every or almost every class in these frequencies; 22%, 23%, 26% and 28%, respectively. Unfortunately, there is no particular traditional teaching method reported in all previous TIMSS waves. The lecture-style method was the only one reported in two waves prior to Bietenbeck (2014)'s study where he identified a small decline from 44% to 43% in proportion of students who reported experiencing this particular teaching methods in every or almost every class. Although this is in line with a tendency in the US to move from traditional towards modern teaching methods, we cannot draw conclusions from such a limited sample. The fact that TIMSS questionnaire contains different teaching methods in each wave prevents to perform analysis over time.

2.1.2 Beyond test scores

Test scores as a measurement of student skills are prevalently used in current literature which exploits the impact of modern and traditional teaching methods of student outcomes. However, one can argue test scores are not the only skill which is developed in schools. Some of the aforementioned papers also investigated impact of teaching methods on cognitive skills (Bietenbeck 2014), socio-emotional skills, such as intrinsic motivation or self-confidence (Korbel & Paulus 2018), and social capital (Algan et al. 2013).

Korbel & Paulus (2018) suggest that implementation of modern methods can contribute to development of students' socio-emotional skills while not having any adverse effect on student test scores in the Czech Republic. In addition, Bietenbeck (2014) splits cognitive skills into three subgroups as knowing, applying, and reasoning. He finds a positive effect of traditional teaching practices on knowing and applying. In regard to reasoning, he identifies modern teaching methods to have a positive effect on reasoning skills. However, these types of skills are outside of the scope of this work due to data availability.

2.2 Across-country differences in student performance

To author's best knowledge she is not aware of any studies investigating why the differences in the effect of a specific teaching practice on test scores exist. Nevertheless, there exists literature which examines international differences in student achievement in general.

Literature suggests an existence of enormous differences in student achievement around the world. Woessmann (2016) proposes that when observing 15-year old students in various countries, the gap in student achievement between top and bottom performing countries is around two standard deviations which equates to approx. 6-8 years of schooling. That is almost the entire length of the obtained education. He suggests that these differences can be explained by educational system (institutional structures), societal setting or cultural factors (Woessmann, 2016).

Moreover, Hanushek & Woessmann (2010) reviewed existing literature and identified three main groups which are explored in previous studies and seem to have explanatory power in explaining the international differences in student achievement. Namely student background, school inputs and institutional structures of educational system. In line with our approach, the third group is of highest relevance for our study. Authors further classify educational system institutions into five subgroups: accountability, autonomy, private school competition, tracking and pre-primary education.

In this particular context, institutions are understood as how educational systems are organized and what national policies are in place at a country level (Hanushek & Woessmann 2010). International student tests such as TIMSS allow researchers to explore variation across countries which is one of the key advantages of standardized tests, as existing variation across countries is more pronounced in these tests (Hanushek & Woessmann 2010).

Looking at accountability as the first subgroup of institutions, it is suggested that curriculum-based external exit exams (CBEEE) serve as an accountability measure (Hanushek

& Woessmann 2010). Not only parents and wider society can easily monitor the educational system through centralized exams, students also tend to be more motivated to perform better and achieve higher results as this is signalling their academic performance to wider society and their future employers (Bishop 1997; Hanushek & Woessmann 2010; Woessmann, 2016).

In general, students in countries with existence of centralized exit exams tend to achieve significantly higher results in standardized tests than students in countries without CBEEE existence, and this effect is observed in international studies as well as region specific studies for Canadian provinces or German states (Bishop 1997; Jürges, Schneider & Büchel 2005; Hanushek & Woessmann 2010). In addition, Fuchs & Woessmann (2007) suggest that use of standardized tests on regular basis by teachers positively contributes to student performance.

There are several challenges related to investigating the effect of the existence of exit exams in countries on student performance. First of all, each country educational system is very unique and differs in various aspects, including differences in the form of the secondary school exit exam itself (Bishop 2005). Thus, one might argue it is impossible to unequivocally assign a country as with or without a centralized exit exam. Bishop (1997) is the first one to consistently advise whether a centralized external exit exam exists in a particular country or not. He defines so called CBEEEs (curriculum-based external exit exam) countries in which subject-specific exams with an externally defined standards are present. According to his definition, these exams are applicable to vast majority of secondary students and have consequences for students. His classification is provided for 39 countries. Second, there is a disagreement among researchers about the definition, and when some countries are reclassified as with an exit exam the positive impact of CBEEE presence on test scores disappears (Huang 2009). Mons (2009) also found that the effect disappears when controlling for GDP per capita.

Besides, research suggests that existence of private schools as another institutional feature plays a role in explaining differences in student performance across countries (Hanushek & Woessmann 2010; Woessmann 2016). Overall, they conclude that higher competition among schools in a given country seems to positively impact student performance. However, they point out challenges connected to the effect of private school competition on student test scores, as low-quality public education leads to an increase in demand for private schools and thus sorting into schools would not be random.

Hanushek & Woessmann (2010) add that student tracking into different school types as an institutional measure may have explanatory power in explaining international differences in student performance. In general, student tracking increases inequalities among students (Hanushek & Woessmann 2010). However, majority of OECD countries tend to perform tracking after 8th grade (Hanushek & Woessmann 2010), which are the students we investigate in this work. Thus, along with a lack of available data, this aspect is not further investigated.

For instance, Korbel & Paulus (2018) accessed a unique dataset for the Czech Republic to examine subject-specific sorting, but such data is not available for a wide portfolio of countries.

Lastly, Hanushek & Woessmann (2010) suggests that organization of pre-primary education: whether its attendance is mandatory, and a different starting age of primary education may explain differences in student achievement across countries. For instance, Schuetz (2009) identifies a positive effect of attending pre-primary education on student performance in 15 countries and this effect is stronger in countries with higher spending on education.

Apart from the educational system, country-level characteristics such as GDP per capita, continental effects or expenditure per student can account for across country variation in student achievement (Hanushek & Woessmann 2010; Woessmann, 2016). However, previous literature suggests that once we control for GDP, the effect of expenditure on education loses its explanatory power (Hanushek & Woessmann 2010). In addition, literature suggests that variation in student achievement across countries can be associated with different Western and Asian values and approach towards education (Hanushek & Woessmann 2010; Bietenbeck 2014).

More specifically, Korbel & Paulus (2018) who investigated the role of modern and traditional teaching methods on students' performance in the Czech Republic suggest that current level of modern practices used in a classroom in a given country may influence the effect of modern teaching methods on student performance. For instance, modern teaching methods can be perceived much newer and unique in countries such as the Czech Republic where their tradition is much shorter.

In addition, research indicates several other characteristics to possible account for the international variation in student achievement in international tests. Cascio, Clark & Gordon (2008) investigated the effect of literacy levels within a particular country and among different age groups in the US. Furthermore, Guiso, Monte, Sapienza & Zingales (2008) examined the role of gender inequality and achievement gap, i.e. cultural attitudes towards women in the society or female economic activity. Also, Korbel & Paulus (2018) suggests gender differences across countries. Lastly, in a wider perspective, institutions may include law, property rights protection, financial institutions, public health systems, labour institutions as well as corruption or governance principles (Chang 2002).

Our work responds to a call for future international research to examinate institutional structures beyond the concepts which were investigated up till today (Hanushek & Woessmann 2010). This is also one of the strong contributions of this study, as there is no previous study investigating across country differences in student performance which would include cultural characteristics. Although it is challenging to capture cultural effects, Hofstede (1991) made an extensive research to quantify cultures and identify differences. Based on a large IBM dataset, he identified four cultural dimensions: power distance, individualism vs. collectivism, masculinity vs. femininity and uncertainty avoidance. All are expressed on a scale from 0 to 100 and the data is available on a country level.

Hofstede's (1991) first cultural dimension power distance expresses how less powerful members of the society cope with unequally distributed power among members of a particular society. In countries with low values of power distance index, people prefer power to be more equally distributed and if it is not, they ask for a valid reasoning of any present inequalities (Hofstede Insights 2020). These are for instance Anglo-Saxon countries such as Australia, England, Scotland or the US. On the other hand, high values of power distance index mean that people strive for hierarchical order and accept hierarchy more easily.

Individualism vs. collectivism dimension expresses the degree to which individuals tend to take care about themselves (and closest family) vs. wider society (Hofstede Insights 2020). High values of this index can be translated as members of the particular society tend to be more individual. In general, western countries such as US tend to be more individualistic with people using "I" a lot compared to Asian countries where "we" is more important. We can expect more individual societies to put emphasis on expressing student own opinion towards a discussed topic. On the other hand, collectivistic societies may value student achievement more as student result represents an entire family.

Third dimension of culture is masculinity vs. femininity which expresses how much society asks for material rewards and heroism (Hofstede Insights 2020). Countries which are more masculine are usually more competitive. On the other hand, feminine cultures are more consensus oriented and care about quality of life and wider cooperation.

Last dimension uncertainty avoidance expresses how people accept uncertainty towards future (Hofstede Insights 2020). Countries with high uncertainty avoidance levels tend to plan future and exhibit more rigid codes of behaviour. Moreover, in these countries, institutions to avoid uncertainty are usually present in these countries and society is strict in deviation from these widely accepted codes of behaviour. This can serve as a proxy in measuring the inclination towards rigidity in a given country. On the other hand, societies with low levels of uncertainty avoidance will let the future happen and have more relaxed attitude (Hofstede Insights 2020).

2.3 Across-country differences in the effect of modern and traditional teaching methods on student performance

Although there are no studies investigating *why* existing research identified differences in the effect of modern and traditional teaching on student performance, there is one study which includes a sign of international comparison.

Bietenbeck (2014) study includes a sign of cross-country comparison as a robustness check for the results for the US. He suggests a positive effect of traditional teaching methods on student test scores while no effect of modern teaching methods. However, the study does not report the results separately for each of these countries. Due to data availability and the nature of his analysis investigating three cognitive skills dimensions, Bietenbeck (2014) limits the analysis to only 9 advanced economies as per International Monetary Fund (IMF) definition which have science taught as an integrated subject (not as biology, chemistry, physics and earth science). His pooled sample containing mixed student data from these 9 countries include Australia, England, Scotland, Israel, US, Hong Kong, Japan, Singapore, South Korea, and Taiwan.

3 Empirical strategy

In this part, we introduce our empirical strategy. The goal of our empirical analysis is to find out the following.

1. Whether the causal effect of modern (or traditional) teaching methods on student's performance differs across countries. (Hypothesis 1)
2. Why do modern (or traditional) teaching methods have impact on test scores in some countries and not in another countries? (Hypothesis 2)

We use student-level data from a large international survey TIMSS (Trends in International Mathematics and Science Study) to approach these questions. We aim at estimating the effect of modern and traditional teaching methods on students' performance, evaluating whether this estimated effect differs across countries and, if it does, finding out which country characteristics are responsible for the differences. This is a typical example of a hierarchical model and this is why we use Hierarchical Linear Modelling (HLM) in our analysis.

Hierarchical levels of grouped data are typical in educational sector where data is organized at various levels: starting from individual students, through classes and schools, to country level. HLM is a complex method which takes into consideration the variance shared at different levels (Woltman et al. 2012). The method simultaneously examines the relationship within and between various levels of the grouped data.

In this thesis the within level analysis corresponds to the within-country relationship between the use of specific teaching methods and students' performance, and the between levels analysis corresponds to the analysis of country characteristics which might affect the within-level estimates. Following this data structure and our research questions, we have decided to apply the slopes-as-outcomes analysis investigating the extent to which the across-countries variation in the effects of modern and traditional teaching styles on students' performance can be explained by country characteristics (Woltman et al. 2012). In other words, we investigate whether the between-countries difference in slopes is related to country characteristics. The full model thus consists of two equations: the within-countries first-stage equation relating student performance to teaching methods, and the between-countries second-stage equation relating this relationship to country characteristics. More specifically, our model is specified as follows:

First stage:

$$TestScore_{ijtc} = \beta_{0c} + \beta_{1c} traditional_{ijtc} + \beta_{2c} modern_{ijtc} + Z_{jtc}\beta_{3c} + a_{ic} + u_{ijtc} \; (1).$$

Second stage:

$$\hat{\beta}_{1c} = \gamma_{10} + country\_characteristics_c \gamma_{11c} + \varepsilon_c \ (2a),$$

$$\hat{\beta}_{2c} = \gamma_{20} + country\_characteristics_c \gamma_{21c} + \nu_c \ (2b),$$

where index *i* stands for a student, *j* for a subject, *t* for a teacher and index *c* for a given country c in equation (1). The outcome variable test score of a student in a specific subject taught by a specific teacher is regressed on traditional $traditional_{ijtc}$ vs. modern $modern_{ijtc}$ teaching methods, teachers and class characteristics Z_{jtc}, and student fixed-effects a_{ic} which accounts for any student characteristics determining the outcome variable. The u_{ijtc} is the error term.

In equation (2a) and (2b), *c* stands for a country. The outcome variable $\hat{\beta}_{2c}$ is our estimated modern teaching practice coefficient from the first stage within-countries analysis. Similarly, $\hat{\beta}_{1c}$ is the estimated traditional teaching coefficient from the first stage within-countries analysis. The dependent variable is regressed on country characteristics $country\_characteristics_c$. The ε_c and ν_c is the error term, respectively.

3.1 First stage

In the first stage, we explore the effect of modern and traditional teaching practices on student performance on a country level for 43 countries for which the relevant data is available. This analysis follows previous literature which, however, cover just few countries level evidence: the Czech Republic (Korbel & Paulus 2018), Israel (Lavy 2016), the Netherlands (Klaveren 2011) or the US (Bietenbeck 2014). Our work goes beyond and explores all countries for which the data is available. We find that there are significant differences among countries in the effect of modern (or traditional) teaching methods on test scores (Hypothesis 1). Thus, our investigation continues and in the second stage, we explore why in some countries modern (or traditional) teaching methods impact test scores while there is no effect identified in other countries (Hypothesis 2).

In the first stage of our analysis, the ideal set up for estimating the effect of teaching practices on students' performance would be the one in which students are randomly assigned to schools and classes with different teaching practices or resources (Krueger 1999). In reality, however, neither students nor teachers are assigned to schools randomly and regressing a variable measuring student's performance on school, teacher and student characteristics overlooks the selection problems.

First, students (or their parents) might choose schools according to their preferences for modern vs. traditional teaching method. For instance, students with high unobservable academic ability may prefer schools that use more modern teaching approach. Thus, not accounting for this sorting pattern would make our estimate biased upwards. Similarly, students with high unobservable academic abilities may influence the teaching style, at least to some degree, which would also lead to the omitted variable bias. Second, teachers may choose to teach at schools which promote their preferred teaching method. For example, teachers with more individual learning style would sort themselves into schools where this teaching style is preferred. This would mean that teaching practices are identified based on unobservable teacher or student characteristics which would again cause a bias.

Following the recent academic literature (Aslam & Kingdon 2011; Schwerdt & Wuppermann 2011; Bietenbeck 2014; Korbel & Paulus 2018), we deal with this selection problem by applying the within-student between-subject approach. Specifically, we rely on the panel data structure of the available dataset and control for student fixed effects. Such a strategy eliminates the bias due to the difference between schools and classes.

To include student fixed effect in the model we need data with at least two observations per student at either different points in time or at the same point in time but for different subjects. TIMSS data exhibits the latter pattern and each student is observed at least twice: in mathematics and science. Following the strategy of Bietenbeck (2014) and Korbel & Paulus (2018), we explore the variation of teaching practices between subjects for each student to estimate the effect of teaching practices on student performance. This strategy accounts for sorting of students to teachers as well as teachers' adaptation of teaching style depending on students' characteristics, because it relies on differences in teaching methods experienced by individual students. In addition, as we observe several students within the same school, the model also allows controlling for school fixed effects. The important assumption behind this estimation strategy is that student and school fixed characteristics have similar effect on our outcome variables across subjects.

The estimation strategy used in the first stage of our analysis can be summarized by the following model.

$$TestScore_{ijtc} = \beta_{0c} + \beta_{1c}traditional_{ijtc} + \beta_{2c}modern_{ijtc} + Z_{jtc}\beta_{3c} + a_{ic} + u_{ijtc} \quad (1)$$

where index *i* stands for a student, *j* for a subject, *t* for a teacher and index *c* for a given country c. The outcome variable test score of a student in a specific subject taught by a specific teacher is regressed on traditional $traditional_{ijtc}$ vs. modern $modern_{ijtc}$ teaching methods, teachers and class characteristics Z_{jtc}, and student fixed-effects a_{ic} which accounts for any student characteristics determining the outcome variable. The u_{ijtc} is the error term.

Parameters of this model are estimated using the fixed effects transformation. It means that averages across subjects from each variable are subtracted from each variable leaving us with demeaned data with respect to subject. This corresponds to the within-student between-subject strategy.

Our empirical strategy of within-student between-subject approach accounts for majority of selection issues, as discussed. For instance, the selection problem of students into schools or classes. However, our approach assumes that students do not sort into teaching methods in schools and classes focused on a specific subject. Hanushek & Woessmann (2010) suggests that majority of countries do not track students into different school types before eight grades. Further assumption we made is that the choice of teaching practices is not correlated with unobservable teacher characteristics. In order to eliminate this concern to minimum, a rich set of teacher and class control variables is controlled for in our regression.

The outcome of the first stage are the estimates of $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$ which are the coefficients capturing the effects of traditional or modern teaching methods on test scores for each country c, respectively. These coefficients will be used as a dependent variable in stage two, which is supposed to estimate the relationship between country characteristics and the impact of modern and traditional teaching methods on students' outcomes.

Before moving on to the second stage, the so-called one-way analysis of variance is performed to check the necessary assumption on variability in the outcome variable in stage two analysis (Woltman et al. 2012). This leads us to testing Hypothesis 1: existence of differences across countries in the relationship between modern (or traditional) teaching methods on students' performance, as suggested by previous scarce literature (Klaveren 2011; Bietenbeck 2014; Lavy 2016; Korbel & Paulus 2018). In the second stage, we take the full list of $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$ estimates from the first stage which enter our second-stage international analysis as a dependent variable being regressed on country characteristics. In this way we test Hypothesis 2.

The first-stage analyses on country level were performed in software STATA. We used command areg for fixed effect regression because order of subjects in which students are observed does not matter (in contrast, the order of time data would be important).

3.1.1 Special data characteristics

The TIMSS dataset used in the first stage analysis requires some special treatment. Student test scores are reported using the multiple imputations methodology, or so-called plausible values. To account for the complex process where not every student answered all questions in the

TIMSS exam booklet, the analyses must be conducted separately for each of the five plausible values reported in the dataset and then the estimates from these separate analyses must be averaged (Martin et al. 2008). This is an advanced and computationally intense method. Thus, STATA software is used to perform the first-stage analyses on country level. This software allows working with multiple imputations for a dependent variable, in our case these are five plausible values, and using the jackknife method to obtain appropriate standard errors for each coefficient in the regression.

Furthermore, clustered standard errors on class level are reported to account for the specific sampling design, as recommended by previous literature (Bietenbeck 2014; Abadie et al. 2017; Korbel & Paulus 2018). Reporting accurate standard errors is a key determinative factor in making inference. This is an additional reason to use STATA software, as it is able to perform our analyses with all aforementioned specifications in contrast to R software.

3.2 Second stage

Following Woltman et al. (2012), our second stage analysis is performed on country (group) level with country characteristics as explanatory variables. Disturbances in the second stage are not completely random but depend on the precision with which the relevant coefficients in our first stage analyses are estimated. The accuracy of the estimation of parameters in the first stage is likely to vary across groups (Hofmann 1997). To rely more on precisely estimated coefficients and give less weight on imprecisely estimated coefficients, the weighted least squares method is used. Following Donald & Lang (2007), we use the estimated standard errors from the first-step regressions as weights in the second-stage analysis. The use of weights also increases efficiency of our second-stage estimates, because homoscedasticity assumption does not hold in HLM models (Woltman et al. 2012).

This leads us to the following models used in our second stage analysis.

$$\hat{\beta}_{1c} = \gamma_{10} + country\_characteristics_c \gamma_{11c} + \varepsilon_c \ (2a)$$

$$\hat{\beta}_{2c} = \gamma_{20} + country\_characteristics_c \gamma_{21c} + \nu_c \ (2b)$$

where c stands for a country. The outcome variable $\hat{\beta}_{2c}$ is our estimated modern teaching practice coefficient from the first stage within-countries analysis. Similarly, $\hat{\beta}_{1c}$ is the estimated traditional teaching coefficient from the first stage within-countries analysis. The dependent variable is regressed on country characteristics $country\_characteristics_c$. The ε_c and ν_c is the error term, respectively.

3.2.1 Bayesian model averaging (BMA)

As mentioned, to the author's best knowledge, there is no existing literature investigating why the differences across countries in the effect of modern or traditional teaching practices on student performance exist. Thus, it is not known in advance what country characteristics should be included in our second-stage model explaining the international differences.

Bayesian model averaging (BMA) method deals with model uncertainty and justifies a particular model choice by "combining predictions from multiple models with weights based on their posterior model probabilities" (Zeugner, 2011; Banner & Higgs 2017, p. 78). Therefore, it serves us as a supporting analysis and a formal check in the problem of variable selection in the second-stage analysis. In particular, BMA method assists when (1) scarce evidence about what explanatory variables should be included in the true model exists and (2) there is a limited number of observations, which is our case (Banner & Higgs 2017).

There exists a BMS package in R software which allows us to perform the BMA method for all our potential explanatory variables. The analysis provides several diagnostics and checks showing probabilities whether a given variable belongs to the final model, i.e. posterior inclusion probability, PIP (Hoeting et al. 1999) and with what sign (i.e. Cond. Pos. Sign - "the posterior probability of a positive coefficient expected value conditional on inclusion") (Zeugner, 2011, p. 4). In addition, it provides information about a model size and an overview of top models with highest posterior model probability (PMP) indicating what proportion of total PMP the top model(s) account for. Lastly, the method offers graphical analysis suggesting a particular model inclusion.

We choose a default option of the Bayesian model averaging in R to examine the ability of the potential country characteristics to predict the differences in the effect of modern and traditional teaching methods on student test scores across countries (Hoeting et al. 1999; Zeugner, 2011). As Bayesian model averaging does not allow including weights, we divide our dependent variable by the estimated standard error from the first-stage country-level analysis. However, due to complexity of BMA method interpretation, we decided to present results of the WLS model as the main ones in this work.

3.3 Hypotheses

To sum up, based on the previous argumentation, the following two hypotheses are tested. Hypothesis 1 is that the causal effect of modern (or traditional) teaching methods on student's performance differs across countries. This is a prerequisite for continuing further and perform investigation of the differences across countries. The second hypothesis is why modern (or traditional) teaching methods have impact on test scores in some countries and not in another countries. In other words, what country characteristics are responsible for the international differences in the effect of modern (or traditional) teaching methods on students' performance. First hypothesis is tested informally for each pair of countries, while our second hypothesis is tested using weighted least squared model, as explained above.

- Hypothesis 1: The causal effect of modern (or traditional) teaching methods on student's performance differ across countries.
- Hypothesis 2: Why do modern (or traditional) teaching methods have impact on test scores in some countries and not in another countries?[1]

[1] In Master's Thesis Proposal, we initially mentioned four hypotheses in total. However, we focus on the effect of modern (and traditional) teaching methods on students's test scores in the end, as such an analysis is already very extensive, complex and resource intensive and presents an important contribution to the existing literature.

4 Data

Based on the empirical strategy, we need two datasets. For the first-stage country level analysis, data on student performance and teaching methods are required. For the second stage international across country analysis, we need data about country characteristics.

4.1 First-stage analysis

4.1.1 Data collection

In the first stage analysis, data from the Trends in International Mathematics and Science Study (TIMSS) which was collected in 2007 among fourth and eighth graders in 59 countries is used. The data is publicly available on the survey's website[2]. The database provides rich information on students' performance, teachers' characteristics as well as questionnaires with student-reported classroom teaching practices.

We use the 2007 TIMSS wave in this work, because it is the last wave including students' reports on teaching practices in a classroom. The particular focus on eight graders is due to two reasons. First, eighth graders student data contain much more detailed information about teaching practices than for fourth graders. Furthermore, majority of these practices in fourth graders questionnaires are subject related, thus not enabling us a consistent between-subject comparison. Second, fourth graders are usually taught by a single teacher in all their subjects in various countries around the world (Bietenbeck 2014; Korbel & Paulus 2018), which would not allow us to use the within-student between-subject empirical strategy as it explores the variation in teaching practices between subjects.

The TIMSS data is collected in a stratified two-stage sampling strategy. First, schools are chosen with a probability proportional to their size and then, one or two classes are randomly sampled within that particular school. This unique format of the data enables to link students to teachers and thus to the classroom practices in each subject. As mentioned earlier, to account for the complex sample design, sampling weights are used whenever we work with individual student data, i.e. our first-stage country-level analyses. Also, special re-sampling approach, the

[2] https://timssandpirls.bc.edu/TIMSS2007/idb_ug.html

so-called jackknife procedure, is used when estimating sampling variance (Joncas 2008; Martin et al. 2008).

Furthermore, the TIMSS data provides special weights for cross-country comparison. These weights allow each country to contribute equal proportion to the international comparison (Foy & Olson 2009; Harmouch et al. 2017). It means that results for any two countries contribute in the same extent upon estimates. Such weights are a transformation of the total weight and result in a weighted sample size of 500 in each country. In contrast, total weight sums up to national population and should be used solely for a single-country analyses. TIMSS team advises the former weights to be used for international comparison (Foy & Olson 2009).

4.1.2 Advantages of data from international standardized tests

International data from standardized tests exhibit unique format with various advantages when performing international comparison. Hanushek & Woessmann (2010) highlights six main advantages of international TIMSS data.

First, such data captures cross-country variation which lies only among countries and cannot be found in solely national data sets, such as differences in institutions. This may include characteristics of a country in respect to teachers, such as competitiveness level to become a teacher, flexibility of teachers in teachers labour market, or teachers' societal status. For instance, it is generally known that teachers have a different social status in Switzerland as compared to the Czech Republic (Varkey Foundation 2018). On institutional level, this can include effect of private schools in a country or existence of centralized exit exams at a country level (Hanushek & Woessmann 2010). Second, any existing variation in institutions and school and population characteristics is more pronounced in international data, increasing the chance to identify impacts when investigating relationships. Third, same data collection process enables an immediate across-country comparison of the identified effect of variables, including the check whether it is country-specific or applies in wider context. Fourth, we can exploit heterogeneity across countries deeper and try to identify factors explaining existence of these differences.

Fifth, the TIMSS data format accounts for the selection problem which majority of national data suffer from. For instance, national data would not allow us to investigate a role of institutions on student performance when comparing pupils attending public and private schools in the same neighbourhood due to unobserved characteristics and selection issues. Thus, aggregation of the information from international surveys at country level allows us to investigate the role of institutions (Hanushek & Woessmann 2010). Moreover, empirical approach in this thesis is chosen carefully to further reduce any bias to minimum in line with previous literature. Lastly, international data allows to capture systemic effect. For example, effect of private schools is further influenced by the competition of nearby schools which national data cannot capture. However, aggregating the effect on a country level and comparing student performance across countries with a certain level of private school competition may capture the role of private education on student results (Hanushek & Woessmann 2010). To conclude, TIMSS data structure is perfect to perform analyses both on national level and to fully exploit cross-country differences.

4.1.3 First-stage dependent variable

Our dependent variable in the first stage analysis, which is performed separately for each country, is student performance in a standardized test. As mentioned, our empirical strategy requires each student to be observed at least twice in order to control for student fixed effects. TIMSS data meets this criterium: students are tested in two up to five subjects in each country: in math and science or in math and biology, chemistry, physics, and earth science, respectively. Following previous studies (Bietenbeck 2014), we standardize test scores within each country to have zero mean and standard deviation of one. This allows for cross-country comparability of first-stage analysis results.

4.1.4 First-stage independent variables: teaching practices index

In order to investigate the effect of teaching practices, aggregate class indexes capturing modern and traditional teaching practices are constructed from student questionnaires in line with previous studies (Bietenbeck 2014; Korbel & Paulus 2018). Students reported how often they experienced a given teaching practice separately for a math and science classes, selecting one of the four following options: ‘in every or almost every lesson’, ‘in about half of the lessons’, in ‘some lessons’ or ‘never’. In some countries, science is taught as an integrated subject whereas in others, science is separated into four individual subjects: biology, chemistry, physics, and earth science (Martin et al. 2008).

We selected three practices which were categorized as traditional teaching methods and another three as modern following Bietenbeck (2014) and Korbel & Paulus (2018). Traditional teaching methods index include memorization of procedures and formulas, listening to lecture-style presentation, and working problems on your own. Whereas, modern teaching methods index include working together in groups, explaining your own answers and relating what you are learning to your daily life.

This particular choice is due to several reasons. In total, there are 16 different practices about which eight-grade students were asked in TIMSS survey, but the remaining methods were either subject-specific, such as conducting an experiment or writing math equations, which does not allow us to construct a consistent index across subjects or the methods were too general, i.e. doing homework or having a test which is not discussed in existing studies (Klaveren 2011; Schwerdt & Wuppermann; 2011; Lavy 2016).

Table 1 introduces the exact questions which students answered in the TIMSS questionnaire about teaching methods they experienced in a class. There is a tiny difference in wording between mathematics and science, but it clearly represents the same activity.

Table 1: Modern and traditional teaching practices indices

Modern teaching practices index: list of individual practices	
Math	We work together in small groups
	We explain our answers
	We relate what we are learning in mathematics to our daily lives
Science	We work in small groups or an experiment or investigation
	We give explanations about what we are studying
	We relate what we are learning in science to our daily lives
Traditional teaching practices index: list of individual practices	
Math	We memorize formulas and procedures
	We listen to the teacher give a lecture-style presentation
	We work problems on our own
Science	We memorize science facts and principles
	We listen to the teacher give a lecture-style presentation
	We work problems on our own

Source: TIMSS 2007 Student Questionnaire (IEA, 2009)

In order to construct the final index, we follow Bietenbeck (2014) and Korbel & Paulus (2018) and take several steps. First, student answers about how often they experienced a given practice in a classroom were rescaled. We assigned a numerical value 0 to the answer ‘never’, 0.25 to ‘in some lessons’, 0.5 to ‘in about half of the lessons’, and 1 to ‘in every or almost every lesson’. Second, we aggregated answers of students at class level and calculate class-level means. In the last step, averages of the three dimensions of traditional and modern teaching methods are computed, resulting in one modern and one traditional index at class level, respectively. Following the explanation of Bietenbeck (2014), we can say that the outcome indexes at class level capture the proportion of lessons in which students were taught using traditional or modern teaching methods, as experienced by students themselves.

4.1.5 First-stage independent variables: teacher and class characteristics

Our first-stage analysis investigates the impact of modern and traditional teaching methods on student performance in each country. It is important to mention that the data structure does not allow to add teacher fixed effect in the first-stage analysis on country level. Thus, we add as many teacher and class characteristics as possible to limit any potential bias due to unobservable teacher characteristics, as TIMSS database contains rich data on teacher and class characteristics (Martin et al. 2008). Missing values in these control variables are replaced with zero values in order to avoid decreasing the number of student observations any further. Descriptive statistics of each country's teacher and class characteristics can be found in Appendix.

The following teacher characteristics data is subtracted from the teacher questionnaires in line with previous literature (Korbel & Paulus 2018; Bietenbeck 2014).

- *Female* – dummy variable which equals one if the teacher is a woman.
- *Age* – teacher's age. The data is split into the following groups: under 25 years old, 25-29, 30-39, 40-49, 50-59, and the age of 60 or more.
- *Experience* – years of experience of teaching all together by the end of the school year (thus, every teacher has a minimum of 1-year experience). We have split the data into the following groups: up to 2 years of experience (base), 3-5 years of experience, and 6 or more years of experience. Previous literature suggests that the effect is largest up to 5 years and then it is relatively constant (Rivkin et al. 2005).
- *Tertiary education* – whether the teacher has a university degree, i.e. ISCED 5 and higher as per international standards. This dummy variable equals one if true.
- *Certificate* – a dummy variable that equals one if teacher holds a teaching licence or certificate.
- *Subject-related development course* – a variable that equals one if the teacher has participated in professional development in mathematics or science pedagogy/instructions.
- *Curriculum improvement course* – a dummy variable that equals one if teacher participated in development in a curriculum.
- *Development in IT integration into subject* – a dummy variable that equals one if teacher participated in a course improving skills related to IT integration into science or mathematics.
- *Course to improve students' critical thinking* – equals one if teacher participated in such a course in last two years.

In addition, we control for the following descriptive class characteristics, also extracted from the TIMSS teachers' questionnaires.

- *Class size* – a discrete variable taking positive values capturing a number of students in a class that took part in TIMSS testing.
- *Teaching time per week* – number of minutes a teacher spends teaching science or mathematics per week.

Furthermore, previous literature suggests that teacher and class control variables do not have an impact on either modern or traditional teaching practices (Schwerdt & Wuppermann 2011; Bietenbeck 2014; Korbel & Paulus 2018). However, as explained above, it is important to include as many control variables as possible to reduce bias due to unobservable teacher and class characteristics to minimum.

4.1.6 First-stage descriptive statistics: modern vs. traditional teaching practices index

In this section, descriptive statistics of teaching practices in classrooms are reported for the entire list of 43 countries which we investigate in this paper. See Table 2. For instance, in Australia we can see that the traditional teaching practices are more commonly used among Australian teachers. On average, teachers spend 50% of their teaching time in a class using traditional teaching methods compared to 47% for modern teaching methods. Data for other countries can be interpreted in the same way.

In general, the minimum proportion of teaching time that teachers spend in a class using *modern* teaching methods is in South Korea: 29%. In contrast, in Jordan teachers spend 65% of their teaching time in a class using modern practices which is the maximum we observe in our sample. On average, teachers spend 52% of their teaching time using modern methods. Looking at *traditional* teaching practices, teachers in Malta spend only 46% of their teaching time using traditional teaching practices which is minimum. Maximum can be found in Jordan again; teachers spend 76% of their teaching time using traditional teaching methods. On average, teachers spend 60% of their teaching time using traditional teaching methods.

Although one can observe a trend towards promotion of modern teaching methods in modern world as discussed earlier, our results suggest that traditional teaching methods remain more prevalent practice in a classroom in the world. This holds for vast majority of the countries, except for Botswana, Cyprus, Ghana and Oman.

In respect to correlation between modern and traditional teaching practices, we observe great differences among countries. The minimum value is 0.074 in Japan while maximum is observed in Georgia with 0.775. On average, the correlation between modern and traditional methods in a classroom is 0.463. In fact, for all countries we observe correlation coefficient to be positive which supports our claim that modern and traditional methods can be viewed as complementary to each other.

It is important to mention that modern and traditional teaching methods do not necessarily dislodge each other and do not need to sum up to one. We believe the two defined teaching methods are complementary to each other, following Bietenbeck (2014) and Korbel & Paulus (2018) approach. Imagine a situation when students are split into small groups in a math class which would be a modern teaching practice while students are told to cooperate and memorize formulas from flashcards which is classified as a traditional teaching practice.

Table 2: Descriptive statistics of teaching practices in a classroom

Country	Teaching practices	Mean	SD	Correlation (modern, traditional)
Armenia	Modern teaching practices	0.51	*0.09*	0.571
	Traditional teaching practices	0.66	*0.10*	
Australia	Modern teaching practices	0.47	*0.09*	0.358
	Traditional teaching practices	0.50	*0.09*	
Bahrain	Modern teaching practices	0.54	*0.09*	0.502
	Traditional teaching practices	0.65	*0.08*	
Bosnia and Herzegovina	Modern teaching practices	0.55	*0.10*	0.603
	Traditional teaching practices	0.62	*0.08*	
Botswana	Modern teaching practices	0.59	*0.06*	0.136
	Traditional teaching practices	0.49	*0.06*	
Chinese Taipei	Modern teaching practices	0.33	*0.07*	0.485
	Traditional teaching practices	0.55	*0.09*	
Colombia	Modern teaching practices	0.61	*0.08*	0.554
	Traditional teaching practices	0.66	*0.07*	
Cyprus	Modern teaching practices	0.53	*0.08*	0.491
	Traditional teaching practices	0.50	*0.07*	
Czech Republic	Modern teaching practices	0.44	*0.08*	0.456
	Traditional teaching practices	0.64	*0.08*	
Egypt	Modern teaching practices	0.62	*0.07*	0.550
	Traditional teaching practices	0.67	*0.07*	
El Salvador	Modern teaching practices	0.58	*0.08*	0.519
	Traditional teaching practices	0.67	*0.07*	
England	Modern teaching practices	0.46	*0.08*	0.410
	Traditional teaching practices	0.55	*0.09*	
Georgia	Modern teaching practices	0.50	*0.11*	0.775
	Traditional teaching practices	0.60	*0.08*	
Ghana	Modern teaching practices	0.65	*0.08*	0.563
	Traditional teaching practices	0.62	*0.09*	
Hong Kong SAR	Modern teaching practices	0.45	*0.07*	0.293
	Traditional teaching practices	0.49	*0.07*	
Hungary	Modern teaching practices	0.45	*0.09*	0.514
	Traditional teaching practices	0.55	*0.09*	
Indonesia	Modern teaching practices	0.47	*0.05*	0.587
	Traditional teaching practices	0.64	*0.04*	
Islamic Rep. of Iran	Modern teaching practices	0.56	*0.10*	0.142
	Traditional teaching practices	0.61	*0.08*	

Israel	Modern teaching practices	0.48	*0.09*	0.378
	Traditional teaching practices	0.65	*0.08*	
Japan	Modern teaching practices	0.52	*0.08*	0.074
	Traditional teaching practices	0.66	*0.07*	
Jordan	Modern teaching practices	0.65	*0.10*	0.640
	Traditional teaching practices	0.76	*0.08*	
Korea	Modern teaching practices	0.29	*0.05*	0.396
	Traditional teaching practices	0.52	*0.06*	
Rep. of Kuwait	Modern teaching practices	0.59	*0.08*	0.499
	Traditional teaching practices	0.69	*0.07*	
Lebanon	Modern teaching practices	0.55	*0.08*	0.395
	Traditional teaching practices	0.62	*0.09*	
Lithuania	Modern teaching practices	0.38	*0.09*	0.477
	Traditional teaching practices	0.63	*0.09*	
Malaysia	Modern teaching practices	0.47	*0.08*	0.629
	Traditional teaching practices	0.56	*0.09*	
Malta	Modern teaching practices	0.45	*0.13*	0.518
	Traditional teaching practices	0.46	*0.12*	
Mongolia	Modern teaching practices	0.50	*0.09*	0.771
	Traditional teaching practices	0.56	*0.08*	
Oman	Modern teaching practices	0.64	*0.09*	0.385
	Traditional teaching practices	0.64	*0.09*	
Palestinian Nat'l Auth.	Modern teaching practices	0.57	*0.09*	0.243
	Traditional teaching practices	0.60	*0.11*	
Qatar	Modern teaching practices	0.59	*0.09*	0.453
	Traditional teaching practices	0.62	*0.07*	
Romania	Modern teaching practices	0.49	*0.12*	0.555
	Traditional teaching practices	0.59	*0.12*	
Russian Federation	Modern teaching practices	0.53	*0.08*	0.523
	Traditional teaching practices	0.70	*0.08*	
Saudi Arabia	Modern teaching practices	0.53	*0.10*	0.583
	Traditional teaching practices	0.60	*0.10*	
Scotland	Modern teaching practices	0.51	*0.08*	0.320
	Traditional teaching practices	0.58	*0.08*	
Serbia	Modern teaching practices	0.46	*0.10*	0.413
	Traditional teaching practices	0.52	*0.09*	
Singapore	Modern teaching practices	0.47	*0.08*	0.523
	Traditional teaching practices	0.53	*0.08*	
Slovenia	Modern teaching practices	0.50	*0.07*	0.452
	Traditional teaching practices	0.66	*0.06*	
Thailand	Modern teaching practices	0.55	*0.07*	0.707
	Traditional teaching practices	0.59	*0.07*	
Tunisia	Modern teaching practices	0.58	*0.08*	0.234
	Traditional teaching practices	0.63	*0.09*	
Turkey	Modern teaching practices	0.55	*0.08*	0.536
	Traditional teaching practices	0.55	*0.08*	
Ukraine	Modern teaching practices	0.55	*0.08*	0.505
	Traditional teaching practices	0.73	*0.08*	
United States	Modern teaching practices	0.53	*0.11*	0.208
	Traditional teaching practices	0.63	*0.09*	

Note: Students weights are used. Standard deviation (SD) in italics. The class-aggregated modern and traditional teaching practices index expresses a proportion of a particular class in which students experience modern or traditional teaching practices.

4.1.7 First-stage sample selection

The 2007 wave of TIMSS testing includes 59 countries in total. We exclude nine countries which did not participate in eight-graders testing, as students usually have the same teacher in all subjects at four-grade (Bietenbeck 2014; Korbel & Paulus 2019) and thus the necessary within-student between-subject variation does not exist. Furthermore, we omit four countries where students could not be uniquely linked to teachers and test scores and where data on teaching practices is not available[3] (Algeria, Morocco, Syrian Arab Republic and Sweden). Besides, one country with missing data on teaching practices on subject level is excluded (Bulgaria). Lastly, we omit two countries where eight-grade students were taught by the same teacher in math and science due to missing the necessary within-student between-subject variation (Italy and Norway). This leaves us with our final sample of 43 countries which we analyse in this work.

On country level, we needed to follow one standardized approach how to treat data in each country in the same way. This is a challenge given unique educational systems and characteristics of individual countries which were not apparent at first glance, especially in regard to science. For instance, there is missing data on teaching practices for a particular science subject in some countries. In Cyprus, we therefore exclude biology. This above is possible due to meeting minimum requirement of two subjects per student which is needed in order to perform our within-student between-subject approach. Besides, Indonesia does not fully fit into either integrated or separate subject educational system but exhibit something in between: biology, physics and integrated science. Thus, we drop integrated science subject due to data availability and focused our analysis on the separate science subjects which is a far most prevalent system in the country.

On country level, we drop student observations with no data on teaching practices or with more than two missing values in teaching practices. In four countries (Armenia, Georgia, Malta and Mongolia), we noticed some students are missing more data on teaching practices in one of the subjects. In that case, we assure each student is observed at least in two subjects. Note that each student observation can be observed in different subjects. For example, due to data quality in Mongolia, each student is observed in exactly two subjects although four subjects

[3] For example, in Algeria, Morocco and Syrian Arab Republic, there are not four separate subjects instead of science, but two merged subjects – biology and earth science as one subject and physics and chemistry as a second subject. There is no data on teaching methods on a subject level. In case of Sweden, the country exhibits something in between separate and integrated subject system and we can find the following subjects in the country: biology, chemistry, physics and integrated science. As integrated science subject is prevalent, there is missing data on teaching methods on subject level as well.

are officially taught in the country. Lastly, in countries such as Malta, several teachers could not have been linked to a particular subject or class.

In our final sample, we include students which were observed twice or more times in one subject. This is especially the case for science classes in integrated science countries. The rationale behind the decision to keep these observations in our study is the following. Dropping these students would reduce our sample and deprive us of useful information. We would need to drop an entire student observation from our sample, as there is no guidance how to choose which class observation to keep and which one to drop when there are multiple observations per one subject. Furthermore, previous literature suggests the results are robust to including students who had several teachers per subject, i.e. attended more classes for one subject in the US (Bietenbeck 2014). In particular, we can see this trend of students attending multiple science classes in e.g., England, Israel, Japan, Scotland, South Korea, and Singapore.

Besides, there are differences across countries in exposition to the same or different classmates in students' lessons. In countries where pupils are usually exposed to the same classmates in all their lessons, such as the Czech Republic (Korbel & Paulus 2018), this fact further lowers any possible effect on the results where students could benefit or suffer from an exposure to different – better or worse – classmates (Hoxby 2000). In contrast, in countries such as the US, students in the same math class do not necessarily attend the same science class. Thus, it implies greater variation in teaching practices in countries where this applies (Bietenbeck 2014). To account for these differences, studies use student sampling weights provided with the TIMSS database throughout their analysis. Our work follows this approach as well.

4.1.8 Final dataset of first-stage analyses

In total, the first stage analysis was performed separately for 43 countries where eight graders were tested. For each of these countries, the estimated effect of modern (traditional) teaching methods on student performance in a given country was obtained. Overview of the final dataset characteristics of the first-stage analysis for each single country is in Table 3. It includes information about the initial number of student observations and the final sample size after excluding students with missing data on teaching practices, as explained earlier. Besides, the Table 3 shows final number of observations, number of students, classes and subjects in each country (i.e. integrated subject countries vs. countries with up to 4 subjects instead of science: biology, chemistry, physics and earth science).

Table 3: Descriptive statistics of first-stage country-level datasets

Country	Number of subjects	Number of observations	Number of students	Initial number of students	% of initial dataset	Number of classes
Armenia	5	19 723	4 016	4 689	85.6%	294
Australia	2	7 902	3 922	4 069	96.4%	238
Bahrain	2	8 220	4 088	4 230	96.6%	201
Bosnia and Herzegovina	5	19 143	3 843	4 220	91.1%	181
Botswana	2	7 950	3 975	4 208	94.5%	151
Chinese Taipei	2	8 131	4 023	4 046	99.4%	153
Colombia	2	9 416	4 708	4 873	96.6%	149
Cyprus	4	16 444	4 121	4 399	93.7%	259
Czech Republic	5	23 480	4 696	4 845	96.9%	212
Egypt	2	12 078	6 039	6 582	91.8%	238
El Salvador	2	7 838	3 919	4 063	96.5%	148
England	2	10 274	3 934	4 025	97.7%	238
Georgia	5	18 261	3 701	4 178	88.6%	184
Ghana	2	10 405	5 008	5 294	94.6%	174
Hong Kong SAR	2	6 959	3 433	3 470	98.9%	120
Hungary	5	19 657	3 919	4 111	95.3%	244
Indonesia	3	10 491	3 863	4 203	91.9%	148
Islamic Rep. of Iran	2	7 818	3 909	3 981	98.2%	208
Israel	2	7 359	3 043	3 294	92.4%	146
Japan	2	10 001	4 273	4 312	99.1%	169
Jordan	2	10 012	5 006	5 251	95.3%	200
Korea	2	9 363	4 226	4 240	99.7%	150
Rep. of Kuwait	2	7 644	3 822	4 091	93.4%	158
Lebanon	4	12 926	3 224	3 786	85.2%	204
Lithuania	5	19 230	3 846	3 991	96.4%	258
Malaysia	2	8 892	4 446	4 466	99.6%	163
Malta	5	12 557	3 637	4 670	77.9%	220
Mongolia	5	8 207	4 103	4 499	92.1%	152
Oman	2	9 100	4 550	4 752	95.7%	158
Palestinian Nat'l Auth.	2	8 328	4 164	4 378	95.1%	153
Qatar	2	13 947	6 808	7 184	94.8%	288
Romania	5	19 722	3 946	4 198	94.0%	266
Russian Federation	5	21 657	4 328	4 472	96.8%	271
Saudi Arabia	2	7 708	3 843	4 243	90.6%	200
Scotland	2	8 939	3 976	4 070	97.7%	244
Serbia	5	18 933	3 790	4 045	93.7%	227
Singapore	2	10 700	4 575	4 599	99.5%	326
Slovenia	4	15 648	3 913	4 043	96.8%	260
Thailand	2	10 730	5 365	5 412	99.1%	150
Tunisia	2	7 809	3 904	4 080	95.7%	169
Turkey	2	8 798	4 399	4 498	97.8%	146
Ukraine	5	20 390	4 078	4 424	92.2%	184
United States	2	14 679	7 199	7 377	97.6%	508

4.2 Second-stage analysis

The main contribution of this work is to extend the existing literature and draw a more general picture of the effect of modern (and traditional) teaching methods on student performance. As explained earlier, our analysis is performed in two stages and in the second stage, we aim to identify country characteristics which could possibly explain the across-country differences. In other words, why modern teaching practices have an effect on student performance in some countries, while there is no effect in other countries. Thus, in the second stage analysis we need data about country characteristics.

4.2.1 Second-stage data collection, dependent and independent variables

In the second stage analysis, our dependent variable is the estimated effect of modern (traditional) teaching methods on student performance in a particular country from the first-stage analysis, this is $\hat{\beta}_{1c}$ and $\hat{\beta}_{2c}$. Our independent variables are country characteristics. Based on existing general scarce literature, author has manually collected data from publicly available databases[4], such as TIMSS additional documentation, the World Bank database, the United Nations database, the International Monetary Fund (IMF) database and the Global Economy (2020) database and compiled a unique dataset. These institutional, educational, cultural and other country characteristics may assist in explaining the international differences in the effect of teaching practices on student achievement, as argued and discussed in Chapter 2.

First, basic country characteristics were extracted from TIMSS Encyclopaedia which is a part of additional documentation available on the website[5]. This includes data on the following.

- *Popul* – size of country's population in millions in a country.
- *Area* – area of a country in square kilometres in a country.
- *Density* – population density, i.e. number of people per square kilometre in a country.
- *Urban* – urban population as percentage of total population in a country.
- *Life* – life expectancy at birth in years in a country.
- *Mortality* – infant mortality rate (per 1,000 live births) in a country.

[4] We take data for the year of 2007, as this is the year when the TIMSS survey took place. In case of England and Scotland, data for the United Kingdom was used. Also, in case of missing data separately for Chinese Tapei (Taiwan), data for entire China region was used in order to avoid reducing our sample of 43 countries any further. Lastly, wherever possible, in case of missing data for 2007 in a publicly available database, data for the preceding or following year(s) was used. The reasoning is that in general, these values do tend to be constant over longer period of time.

[5] https://timssandpirls.bc.edu/TIMSS2007/idb_ug.html

Furthermore, author herself created the following variables from the information available in the TIMSS Encyclopaedia. In particular, the TIMSS documentation contains a section explaining educational system for each country which serves as basis for creating the latter two variables. Any further country characteristics could not be extracted from the Encyclopaedia due to missing data.

- *Nursery* – a dummy which equals one when pre-primary schooling is compulsory in a given country.
- *x6less*[6] – a dummy which equals one if a starting age of a compulsory primary schooling is at the age of 6 or less in a given country.
- *Emphasis_m_s* – equals one if there is a special emphasis on teaching math or science in a country, i.e. a government program to promote teaching math or science subjects in schools in a given country.
- *Exit_exam* – equals one if a secondary curriculum-based external exit exam (CBEEE) exists in a country. The original CBEEE index was defined by Bishop (1997) for 39 countries signalling whether a particular country has a centralized secondary school leaving exam in place. Unfortunately, the information is not available for vast majority of countries from our sample. Thus, in line with Bishop (1997) I extended the existing CBEEE database and assigned value one to countries where national centralized regular secondary school exam exists.

The World Bank (2020) database served as another source of country characteristics data.

- *GDP_PPP*[7] – gross domestic product (GDP) per capita, measured in purchasing power parity (PPP) in current international dollar.
- *Gov_exp_on_edu* – government expenditure on education as percentage of GDP in a country. It takes values between 0 and 100.
- *Female_employment* – female labour force as percentage of total labour force in a country. It takes values between 0 and 100.
- *Prop_prim_private* – percentage of students enrolled into private institutions at primary level of education. The variable takes values between 0 and 100.

---

[6] In case of missing data in TIMSS 2007 Encyclopaedia for any of the previously mentioned country characteristics (3 cases – mortality, nursery, x6less), we used data from the TIMSS 2011 Encyclopaedia available from its website: https://timssandpirls.bc.edu/TIMSS2007/encyclopedia.html. The reasoning is that these characteristics are rather stable over time and we aim to avoid decreasing our sample of 43 countries any further.

[7] For variables such as GDP, there exists various alternative measures of the same country characteristics. In that case, we chose variables with the least missing values in order to avoid reducing our final dataset of 43 countries any further.

- *Migrant_stock* – international migrant stock as percentage of population. It takes values between 0 and 100. The data is collected every five years, we take 2005 values.
- *Oil* – oil rents as percentage of gross domestic product (GDP) in a given country.

In addition, data on the following country characteristics was collected.

- *Index_mod* – a current total average level of modern teaching practices used in a classroom in a country. Author calculated this variable herself from the TIMSS data which was used in our first-stage country level analysis. This variable takes values between zero and one, expressing a proportion of a class where modern teaching methods are used.
- *Mean_y_school* – mean years of schooling of population over 25 years old in a given country (United Nations database 2020).
- *Advanced* – a dummy variable which equals one if a country was classified as an advanced economy in by the International Monetary Fund (IMF 2007).
- *New_asian_values* – whether a country can be classified as with "Asian" values. Based on the GLOBE study (Mensah & Chen 2013), I created a dummy variable which equals one if a country is part of the Confucian or South-East Asian cultural cluster. Continents dummies were not included due to our limited sample size and because Asia also includes the Middle East countries which can be argued are very different.
- *Post_communist* – a dummy variable which equals one if a country can be classified as a post-communist country (CPDS 2008)

Furthermore, the Global Economy (2020) database was accessed and the following information about countries was collected.

- *Pol_stability_i* – political stability index. It takes values from -2.5 up to 2.5. The higher the value, the more political stability in a given country.
- *Corrup* – control of corruption in a country. It takes values from -2.5 up to 2.5 where the higher value, the stronger control of corruption.
- *Agri* – employment in agriculture as percent of total employment in a country. It takes values from 0 up to 100.

Lastly, we collected data about culture (cultural dimensions) in a given country from a publicly available database Hofstede Insights (2020). More details about what a particular index means can be found in Chapter 2.

- *PD* – power distance index takes values between 0 up to 100. Higher values correspond to a country where people prefer power to be more equally distributed.

- *INDIV* – individualism vs. collectivism. It takes values between 0 up to 100 and higher values mean that a country is more individualistic.
- *MAS* – masculinity vs. femininity expresses how much society asks for material rewards and heroism. It takes value between 0 up to 100, where higher values correspond to a more masculine country.
- *UNA* – uncertainty avoidance index takes values between 0 up to 100. The higher the value, the more people avoid uncertainty in a country[8].

In particular, we believe that cultural dimensions may play an important role in explaining the differences in the effect of a particular teaching method on student performance. Existing literature presents ambiguous findings for various variables and calls for research of international differences beyond existing concepts which were investigated up till today (Hanushek & Woessmann 2010). To my best knowledge, there are no studies investigating culture in explaining international differences in the effect of modern teaching methods on student performance.

Let us provide further intuition for this particular choice of country characteristics mentioned above. For instance, it could be that an introduction of new modern teaching methods in more rigid and structured countries where people believe norms should be followed (*UNA*, *uncertainty avoidance index*) can easily draw students' attention and make it easier for students to learn, which in turn could improve their performance. One could argue that a similar trend could be observed in more educated (*mean_y_school*) or countries with higher GDP per capita (*GDP_PPP*) where teachers need to come up with more modern teaching methods in order to attract students, who are used to a lot of distraction, and in order to improve their performance in standardized tests. One could also argue that once the modern teaching methods are used too frequently in a country (*index_mod*, *current level of the use of modern teaching practices*), the *traditional* teaching methods might in fact improve student performance. Furthermore, the existence of centralized exams (*exit_exam*) as an accountability measure and proxy for importance given to education in a given country could indicate that students in these countries are more self-motivated and thus a particular teaching method may have no effect on student performance. Similar rationale applies for the remaining country characteristics.

[8] To avoid reducing number of countries with available data in our second-stage international analysis any further, data on cultural dimensions for Bahrain were obtained from Steers, R. M., L. Nardon & C. J. Snaches-Runde (2013): *Management across cultures: Developing global competencies.* Cambridge University Press. ISBN-13: 978-1107645912

4.2.2 Second-stage final dataset

Our manually collected rich dataset with various country characteristics served as a starting point in the second-stage analysis. At first, author randomly tried which variables have explanatory power and should be included in our final model specification in line with previous literature and economic intuition. However, we were facing several challenges. First, the existing literature does not provide a strong evidence which variables are important and should be included in our true model. Second, we have a limited number of observations in the second-stage analysis – only 43 countries and thus it would be impossible to perform the analysis for all possible explanatory variables due to insufficient number of degrees of freedom. Thus, we followed **Bayesian model averaging** (BMA) method in order to tackle the above-mentioned challenges.

First, BMA assists in situations where many potential explanatory variables which could explain the differences across countries in the effect of modern (traditional) teaching method on student performance exist, but the literature does not provide strong evidence which ones to choose and how important they are (Zeugner 2011). Second, BMA assists when it is not only inefficient to perform the analysis with all potential variables, but also impossible due to a limited number of observations, which is our case (Zeugner 2011). To conclude, BMA takes into consideration all possible model specifications and provides a weighted average of all the possible models and it justifies a particular model choice (Zeugner 2011; Banner & Higgs 2017).

In the second-stage analysis, the following explanatory variables are included in our preferred model specification, following our intuition, previous literature and supported by Bayesian model averaging (BMA) method which tackles model uncertainty. Descriptive statistics overview of the variables included in the preferred model specification can be found in Table 4 below. For descriptive statistics of all 28 potential variables which were considered in BMA method when investigating what could explain the international differences in the role of modern (and traditional) teaching methods on students' performance see Appendix.

Table 4: Descriptive statistics of variables in the preferred model specification in the second-stage international analysis

Independent variable	Min	Max	Mean	SD
Uncertainty avoidance index	8	96	70.395	21.737
Purchasing power parity (GDP per capita, PPP)	2501	122884	25359	24152
Mean educational achievement (in years)	5.648	13.256	9.757	2.331

In addition, correlation coefficients of the explanatory variables in our preferred model specification in the second-stage analysis investigating international differences among countries in the effect of modern (and traditional) teaching practices on students' test scores are presented in Table 5 below. The correlation is very small which is appreciated. Correlation table for all 28 potential variables which were considered in the BMA can be found in Appendix.

Table 5: Correlation between our explanatory variables in the preferred model specification of the second-stage analysis investigating international differences among countries

Independent variable	Uncertainty avoidance index	Purchasing power parity (GDP per capita, PPP)	Mean educational achievement (in years)
Uncertainty avoidance index	1.000		
Purchasing power parity (GDP per capita, PPP)	-0.236	1.000	
Mean educational achievement (in years)	-0.161	0.161	1.000

5 Presentation and discussion of results

5.1 First-stage analysis: country-level results

This thesis investigates the effect of modern and traditional teaching practices on student performance across countries and why these differences across countries exist. In the first stage, we estimated the effect of modern and traditional teaching practices on students' test scores in 43 countries with available TIMSS data. Each country-level analysis was performed using the within-student across-subject design.

This work extends the portfolio of countries for which the country-level evidence about the effect of the modern (traditional) teaching methods on student performance exists. For instance, no effect of either teaching method on student test scores was found in the Czech Republic and in the Netherlands (Korbel & Paulus 2018; Klaveren 2011), positive effect of traditional teaching methods was found in the US (Bietenbeck 2014) and positive effect of both modern and traditional teaching methods was found in Israel (Lavy 2016).

It is imperative for policy makers to understand if these effects are country-specific and apply solely for that particular country, or if the outcomes can be generalized and policy makers in similar countries could implement such policies as well. Or whether each country is unique and no one-fit-all approach exists. Thus, we perform second-stage analysis which examines the international differences and aims to identify patterns across countries, using the weighted least-squares estimation method (WLS). In this section, our first-stage country-level findings are discussed and compared to previous studies.

Results for 43 countries from the first-stage analysis can be found in Table 6. The table reports estimated coefficients corresponding to the effects of modern and traditional teaching practices on students' performance in each analysed country. Significance indicators are included, showing whether the estimated effect of modern or traditional teaching practices is statistically significant in a given country.

Table 6: Results from first-stage country-level analysis

Country	Estimated effect of modern teaching index	*SE*		90% confidence interval	Estimated effect of traditional teaching index	*SE*		90% confidence interval
Armenia	0.221	*0.214*		(-0.131;0.573)	0.091	*0.159*		(-0.171;0.353)
Australia	0.078	*0.159*		(-0.184;0.339)	0.421	*0.213*	*	(0.071;0.771)
Bahrain	0.629	*0.185*	***	(0.325;0.934)	-0.132	*0.273*		(-0.581;0.317)
Bosnia and Herzegovina	0.442	*0.118*	***	(0.248;0.636)	0.316	*0.207*		(-0.025;0.656)
Bostwana	-0.041	*0.204*		(-0.376;0.294)	0.191	*0.169*		(-0.088;0.469)
Chinese Tapei	0.314	*0.236*		(-0.074;0.701)	0.354	*0.185*	*	(0.05;0.658)
Colombia	0.332	*0.130*	**	(0.118;0.547)	0.422	*0.191*	**	(0.108;0.736)
Cyprus	0.195	*0.122*		(-0.005;0.396)	0.211	*0.140*		(-0.02;0.441)
Czech Republic	0.190	*0.159*		(-0.071;0.45)	0.098	*0.166*		(-0.174;0.371)
Egypt	0.126	*0.138*		(-0.101;0.353)	0.117	*0.148*		(-0.125;0.36)
El Salvador	0.443	*0.273*		(-0.006;0.892)	0.082	*0.346*		(-0.487;0.651)
England	0.213	*0.153*		(-0.039;0.464)	0.510	*0.165*	***	(0.238;0.782)
Georgia	-0.001	*0.183*		(-0.303;0.301)	0.320	*0.211*		(-0.028;0.667)
Ghana	0.155	*0.243*		(-0.245;0.554)	0.063	*0.218*		(-0.296;0.421)
Hong Kong SAR	-0.001	*0.193*		(-0.317;0.316)	0.601	*0.304*	*	(0.101;1.102)
Hungary	0.163	*0.109*		(-0.015;0.342)	0.204	*0.126*		(-0.004;0.412)
Indonesia	0.020	*0.481*		(-0.771;0.812)	0.119	*0.313*		(-0.396;0.633)
Islamic Rep. of Iran	0.103	*0.134*		(-0.118;0.324)	-0.115	*0.169*		(-0.393;0.163)
Israel	0.108	*0.189*		(-0.203;0.419)	0.561	*0.174*	***	(0.275;0.848)
Japan	0.400	*0.130*	***	(0.187;0.614)	0.256	*0.155*		(0.001;0.511)
Jordan	0.498	*0.158*	***	(0.239;0.757)	0.011	*0.177*		(-0.28;0.303)
Korea	0.204	*0.326*		(-0.333;0.741)	0.671	*0.276*	**	(0.216;1.125)
Rep. of Kuwait	1.196	*0.123*	***	(0.994;1.399)	-0.271	*0.348*		(-0.843;0.301)
Lebanon	-0.031	*0.219*		(-0.392;0.33)	0.267	*0.156*	*	(0.01;0.525)
Lithuania	0.144	*0.090*		(-0.004;0.292)	0.270	*0.115*	**	(0.081;0.459)
Malaysia	-0.712	*0.244*	***	(-1.113;-0.312)	1.124	*0.535*	**	(0.244;2.005)
Malta	0.086	*0.091*		(-0.063;0.235)	0.007	*0.097*		(-0.152;0.166)
Mongolia	-0.063	*0.529*		(-0.933;0.807)	0.562	*0.575*		(-0.384;1.508)
Oman	-0.199	*0.215*		(-0.553;0.155)	-0.111	*0.259*		(-0.538;0.316)
Palestinian Nat'l Auth.	0.128	*0.144*		(-0.109;0.364)	0.171	*0.203*		(-0.162;0.505)
Qatar	1.065	*0.119*	***	(0.87;1.26)	-1.836	*0.123*	***	(-2.038;-1.634)
Romania	0.119	*0.155*		(-0.136;0.374)	0.242	*0.153*		(-0.009;0.494)
Russian Federation	0.467	*0.150*	***	(0.22;0.713)	-0.153	*0.129*		(-0.365;0.059)
Saudi Arabia	0.223	*0.243*		(-0.177;0.623)	-0.457	*0.338*		(-1.013;0.099)
Scotland	-0.193	*0.131*		(-0.409;0.022)	0.868	*0.171*	***	(0.587;1.148)
Serbia	0.243	*0.102*	**	(0.076;0.411)	0.182	*0.158*		(-0.078;0.442)
Singapore	0.182	*0.101*	*	(0.016;0.348)	0.510	*0.157*	***	(0.252;0.769)
Slovenia	0.145	*0.154*		(-0.109;0.398)	0.335	*0.212*		(-0.014;0.685)
Thailand	0.284	*0.221*		(-0.079;0.647)	0.718	*0.239*	***	(0.325;1.111)
Tunisia	0.520	*0.202*	**	(0.187;0.853)	0.239	*0.392*		(-0.406;0.885)
Turkey	0.115	*0.168*		(-0.162;0.392)	0.096	*0.174*		(-0.19;0.382)
Ukraine	0.423	*0.121*	***	(0.223;0.622)	0.103	*0.167*		(-0.171;0.377)
United States	0.091	*0.115*		(-0.098;0.28)	0.462	*0.109*	***	(0.282;0.641)

Note: Clustered standard errors (SE) are in italics. Test scores is our dependent variable. In each country-level first-stage analysis, test scores are estimated from 5 plausible values. Each country-level regression controls for teacher and class characteristics and subject dummies. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The interpretation of the estimated coefficients from our first-stage country-level models is explained on the example of Bahrain. As mentioned, the modern teaching index corresponds to the share of teaching time in a classroom devoted to modern teaching methods. In Bahrain, students experience modern teaching methods to be used 54% of the teaching time in a class (see Table 2 in Chapter 4). The estimated coefficient of the impact of modern teaching methods on student performance is 0.629 (see Table 6 above). This means that when a student experiences an increase in the exposure to the modern teaching methods by 0.1 (i.e. 10 percentage points from a current level of 54% to 64%), it results in 0.1 x 0.629 = 0.0629 points increase in the student performance. As student test scores were standardized to have zero mean and standard deviation of one in the full sample within a country, we can interpret this 0.0629 points increase as a 6.29% of standard deviation increase in the mean student performance. Furthermore, a 10-percentage point (0.1) change in the use of modern teaching methods can be translated into 4:30 minutes in a 45-minute class. The same interpretation applies to the estimated coefficient of the impact of traditional teaching methods on student performance, holding modern teaching methods constant.

Our results for the Czech Republic and the US are consistent with the previous literature. In case of the Czech Republic, we found no effect of modern or traditional teaching methods on student performance. In the case of the United States, our results suggest positive and significant effect of traditional teaching practices on student achievement while no effect of modern teaching methods on student test scores. This is in line with previous studies for the Czech Republic (Korbel & Paulus 2018) and the United States (Schwerdt & Wuppermann 2011; Bietenbeck 2014), respectively. In case of Bietenbeck (2014) and Korbel & Paulus (2018) studies, this should not be a surprise, as the same strategy was followed, and the same data was used. Our estimates may slightly differ due to a different use of weights for international comparison and the fact that the dataset was treated a bit differently. The promotion of modern teaching practices in the US does not seem to be validated by our findings when looking solely at student performance. However, Bietenbeck's (2014) results suggests that modern teaching practices can contribute to development of other student skills. Unfortunately, due to data availability, we could not perform such analysis for all 43 countries.

Furthermore, our findings show that traditional teaching methods have positive effect on student performance while modern practices have no effect on student results in Australia, England, Scotland, Hong Kong, Korea and Taiwan. Previous research did not investigate these countries separately, but there exists one study (Bietenbeck 2014) which pooled a sample of 9 advanced economies and found a positive impact of traditional teaching methods on student performance. In case of Australia, England, Scotland, Hong Kong, Korea and Taiwan our results correspond with Bietenbeck (2014). However, we found rather opposite results in case Japan and Singapore, which were also included in the Bietenbeck's sample. In case of Japan, we found traditional teaching practices having no effect on students' test scores while modern teaching methods have a positive effect on students' test scores. In Singapore, our results indicate that the both modern teaching practices and traditional teaching practices have positive effect on students' test scores, but the effect is stronger for traditional teaching methods. Thus, our results suggest that the pattern found by Bietenbeck (2014) cannot be generalized to all advanced economies.

Other studies on this topic followed the same empirical strategy but defined teaching practices differently. In case of Israel, our results suggest a positive effect of traditional teaching methods on student performance while no effect of modern teaching methods. This is partially in line with the findings of Lavy (2016), who used national panel data on students who were in fifth grade in 2002 and eighth grade in 2005 and tested in four subjects. Lavy (2016) reports a positive impact of *both* modern and traditional teaching practices on student performance in Israel. His results suggest the effect is larger in case of traditional teaching methods and specific to students from low socioeconomic background in Israel. In contrast, our results reveal a positive effect of traditional teaching methods on student performance but no effect of modern teaching methods. Israel's reform promoting reduction in the use of traditional teaching methods does not seem to be fully validated by our results. Future research should clarify the role of modern teaching methods in development of other students' skills.

Our results can be grouped as follows. Comparing our results from individual country-level analysis, we identified the following effects of modern and traditional teaching methods on student performance. First, in some countries (10), there was found a positive significant effect of *modern* teaching practices on student performance: Bahrain, Bosnia and Herzegovina, Japan, Jordan, Kuwait, Qatar, Russian Federation, Serbia, Tunisia and Ukraine. On the other hand, our results suggest that *traditional* teaching methods positively impact student performance in the following countries (12): Australia, Taiwan, England, Hong Kong, Israel, Korea, Lebanon, Lithuania, Malaysia, Scotland, Thailand and the US. In majority of the aforementioned countries except for Malaysia and Qatar, we found no effect of the other (traditional or modern) teaching method on student performance, respectively.

In Qatar, our results suggest that modern teaching methods have a positive effect on students' test scores while traditional teaching methods negatively impacts student performance. In Malaysia, we found a positive effect of traditional teaching methods on student performance and negative effect of modern teaching methods on students' test scores.

Furthermore, we have identified both modern and traditional teaching methods to positively impact student performance in two countries: Colombia and Singapore. However, in both cases the effect of traditional teaching methods on students' test scores is stronger. We may interpret that higher use of modern and traditional teaching practices will have positive effect on student performance, but the effect is stronger for traditional practices.

Last but not least, in almost half of the countries (19) in our sample, we found neither modern nor traditional teaching practices to have any effect on student performance. This is the case of Armenia, Botswana, Cyprus, Czech Republic, Egypt, El Salvador, Georgia, Ghana, Hungary, Indonesia, Iran, Malta, Mongolia, Oman, Palestinian Nat'l Auth., Romania, Saudi Arabia, Slovenia and Turkey. This is in line with some previous studies which found no effect of both modern and traditional teaching practices on student performance in the Netherlands and in the Czech Republic (Klaveren 2011; Korbel & Paulus 2018). Unfortunately, the Netherlands did not participate in our wave of TIMSS testing for eight-graders, what makes it impossible to validate Klaveren's results in our study. One possible explanation why no effect was identified is a small variation within student observations. Our model explores the within-student between-subject variation, and it could be that there is guidance for all teachers within a given school or region in regard to teaching methods.

To summarize, our findings suggest various impacts of either modern or traditional teaching methods on student test scores: positive effect, no effect or negative effect. Translating this into policy recommendations, in countries where positive effect of modern teaching practices was identified, these countries are suggested to rather promote student-centred modern teaching methods, such as working in groups or relating what are students taught to daily lives, in order to improve student test results in international tests.

On the other hand, in countries where positive effect of traditional teaching methods on student performance was found, it can be recommended to promote the use of traditional teaching methods, such as passive lecture-style teaching or memorization of formulas, in order to improve student performance in standardized tests. The recommendations are even more pronounced in countries where the other teaching method (modern or traditional) have negative effect of students' performance. For the remaining countries, no effect of both modern and traditional teaching methods on student performance was identified.

5.2 Testing the existence of across country differences

As discussed, our results suggest there exists variation in the effect of modern and traditional teaching methods on student performance across countries. This is our first hypothesis that the causal effect of modern (or traditional) teaching methods on student test scores differs across countries. It is a prerequisite for our further investigation of the differences across countries.

The first hypothesis was tested in the following way. Looking at modern teaching methods separately, for each pair of countries, the author calculated whether the respective coefficients are significantly different from each other.[9] Table 6 above reports 90% confidence intervals for the modern teaching index in each country. Differences between the estimated coefficients of the effect of *traditional* teaching methods across countries are tested in the same way and also presented in the Table 6.

The following patterns were identified. Generally, our results suggest significant differences in the effect of modern teaching methods on students' test scores across countries in our sample. First, Europe is quite homogenous with observed no effect of modern teaching methods on students' performance. Second, Asian countries exhibit rather higher estimated coefficients than European countries. Similarly, Middle East countries tend to have higher estimated effect of modern teaching practices on students' test scores.

Last but not least, there exists several countries which significantly differ from the rest in our sample. The estimated effect of modern teaching practices on test scores in Qatar and Kuwait is significantly larger than in all countries, while not being distinct from each other. Similarly, the effect is significantly larger in Bahrain, Bosnia and Herzegovina and Jordan than in majority of the countries. In contrast, in Malaysia the effect is significantly lower than in vast majority of the countries from our sample.

Looking at *traditional* teaching practices, we can also observe significant differences in the effect of traditional teaching practices on student performance across countries. For instance, the effect of traditional teaching practices on test scores in Qatar is significantly lower than in all countries in our sample. Also, the estimated effect found in Scotland and England is significantly higher than in vast majority of countries and in many Middle East countries, respectively.

To conclude, significant differences in the effect of modern and traditional teaching practices on student performance were identified between individual countries. Descriptive statistics of the estimated effect of modern and traditional teaching practices on student performance from first-stage country-level analysis are presented in Table 7 below.

---

[9] For our purposes, informal testing of the Hypothesis 1 is fine. In order to test the Hypothesis 1 formally, it would be necessary to pool the data for all countries together and interact the modern and traditional teaching methods indexes with country dummies. Then, we would test whether the coefficients corresponding to these interactions differ or are the same for any pair of countries. However, this is impossible to do in practice as our datasets slightly differ between countries in our sample.

Table 7: Descriptive statistics of dependent variables in the second-stage international analyses

Dependent variable	Min	Max	Mean	SD
Estimated effect of *modern* teaching practices on student performance from first-stage country-level analyses	-0.712	1.196	0.210	0.307
Estimated effect of *traditional* teaching practices on student performance from first-stage country-level analyses	-1.836	1.124	0.202	0.438

5.3 Second-stage analysis: investigating across country differences

This work contributes to scarce literature investigating the effect of modern and traditional teaching methods on student performance. We are not selective to a particular country but perform the analysis separately for all 43 countries for which data is available in order to present a more consistent evidence. As discussed in the previous section, great variation in the effect of modern (or traditional) teaching methods is observed across countries, which motivates us to investigate further why do modern (or traditional) teaching methods have different impact on student performance across countries (Hypothesis 2). Our aim is to identify patterns from the country-level analysis which could possibly explain the international differences.

In the second-stage analysis, the independent variable is the estimated effect of modern (traditional) teaching practices on test scores in the given country from the first-stage analysis. Our dependent variables in the second-stage analysis are country characteristics, including institutional, educational and cultural characteristics. Standard errors of the estimated effects of teaching methods on student performance are used as weights to account for varying precision of country-level estimates.

5.3.1 Across-country differences: modern teaching methods

The second-stage model investigates the international differences in the effect of modern teaching methods on student performance. The particular choice of independent variables is based on previous literature investigating the differences in the role of modern and traditional teaching methods on student performance in general not focusing on a particular teaching method. Furthermore, economic intuition assisted in identification of possible patterns and what could possibly explain the differences in the role of a particular teaching method on student performance across countries. The author collected unique dataset containing 28 variables on country characteristics, including educational system institutions and cultural aspects, as discussed in previous chapters.

As explained, we were facing several challenges: existing literature did not provide strong evidence of what country characteristics are the most relevant ones and our number of observations (number of countries) is limited. Thus, Bayesian model averaging (BMA) method, which deals with model uncertainty and tackles the abovementioned issues, assisted us and supported our preferred model choice.

BMA unambiguously suggest the following two explanatory variables to be included in the true model: cultural dimension uncertainty avoidance (*UNA*) and gross domestic product per capita, measured in PPP (*GDP_PPP*). Furthermore, it suggests mean educational achievement in years in a given country (*mean_y_school*) to be included in the true model with a posterior inclusion probability over 30%. More detailed about BMA results can be found in the next session.

Table 8: Results from second-stage international analysis explaining international differences in the estimated impact of modern teaching practices on student performance

	Estimated effect of modern teaching index			
Independent variable	**(1)**		**(2)**	
Uncertainty avoidance index	0.008	***	0.008	***
	0.002		*0.002*	
Purchasing power parity (GDP per capita, PPP)	0.00001	***	0.00001	***
	0.00000		*0.00000*	
Mean educational achievement (years)			-0.048	***
			0.015	
No. of observations	38		38	
R squared	0.466		0.594	

Note: Standard errors (SE) are reported in italics. Standard errors of the estimated effect of the modern teaching practices on student performance from the first-stage country-level analysis are used as weights. * p < 0.1, ** p < 0.05, *** p < 0.01

Our preferred model specification is based on BMA method and contains the aforementioned three variables: cultural dimension uncertainty avoidance (*UNA*), gross domestic product per capita, measured in PPP (*GDP_PPP*) and mean educational achievement in years (*mean_y_school*). Estimation results from the two best models are presented in Table 8.

The interpretation of the coefficients is not so straightforward. Due to a complex interpretation of the coefficients, our contribution is rather that our results suggest which country characteristics may explain the differences across countries in the effect of the use of modern teaching methods on student test scores and the direction of the effect. It is important to bear in mind that our dependent variable is the estimated effect of modern teaching methods on student test scores in standardized tests obtained in the first-stage country-level analysis. We calculated that it has a mean of 0.210 and standard deviation of 0.307 (see Table 7 in Chapter 5).

Looking at the second-stage analysis, the coefficients can be interpreted as follows. For instance, let's look at uncertainty avoidance index which is cultural characteristics summarized using a 100 points scale. A 20-point increase in the uncertainty avoidance index (e.g. from 50 to 70), i.e. about one standard deviation increase in the index, accounts for a 0.008 x 20 = 0.16 points increase in the estimated effect of modern teaching on student performance in a country. This corresponds to about half of the standard deviation (SD) of the estimated first-stage coefficient, what is a sizable effect.

Similar interpretation applies for the remaining coefficients. In case of the mean years of schooling, a 2-point increase in the mean length of schooling (e.g. from 8 to 10 years), which is about one standard deviation, accounts for (-0.048) x 2 = -0.096 points change in the estimated effect of modern teaching practices on student test scores. This corresponds to about negative one third of the standard deviation of the estimated first-stage effect. Lastly, in case of purchasing power parity (PPP), a 24,000-point increase in the PPP (e.g. from 10,000 to 34,000), i.e. about one standard deviation increase in the variable, accounts for a 0.00001 x 24,000 = 0.24 points increase in the estimated effect of modern teaching on student performance. This effect corresponds to almost one standard deviation of the estimated first-stage coefficient. These are sizeable effects.

Our results suggest that uncertainty avoidance as a cultural dimension is positively related to the estimated effect of modern teaching methods on students' test scores in standardized tests from the first-stage country-level analysis. As explained, uncertainty avoidance expresses how people accept uncertainty towards future (Hofstede Insights 2020). Countries with high uncertainty avoidance levels exhibit more rigid codes of behaviour and have strong institutions in place (Hofstede Insights 2020). Such societies typically do not accept deviation from accepted codes of behaviour. In other words, it indicates institutional and behavioural rigidity in a given country.

Thus, our findings indicate that in more rigid countries there tends to be a stronger positive effect of modern teaching methods on student performance. Using common sense, this finding is intuitive. In countries where people follow more rigid codes of behaviour, modern teaching methods such as working in groups can be perceived as a newer method and easily draw students' attention and thus positively influence their test scores.

Furthermore, our results suggest that purchasing power parity positively influences the effect of modern teaching practices on student performance. However, this variable serves more as a control variable and the effect seems to be driven by rich Middle East countries, most specifically Qatar and Kuwait.

Lastly, we found that level of education, i.e. the mean years of schooling in a given country, is negatively related to the estimated effect of modern teaching methods on student test scores in standardized tests. It can be interpreted that the estimated effect of modern teaching methods on test scores is higher in less educated economies. This effect is also intuitive. In less educated countries, modern teaching methods, such as working in groups or relating what are students taught to their daily lives, may be viewed as a more unique teaching practice and easily draw students' attention and therefore positively influence students' performance.

To conclude, we found that one of cultural dimensions plays an important role in explaining the across country differences in the role of modern teaching methods on students' test scores. More specifically, uncertainty avoidance as a cultural dimension is positively related to the effect of modern teaching practices on student performance. Modern teaching methods might be more unique in more rigid countries (i.e. higher values of uncertainty avoidance index) and thus their effect on student performance can be more pronounced. In countries with high uncertainty avoidance levels people prefer clear order, strong institutions and incline towards rigidity. For instance, in our sample, the following countries exhibit high levels of uncertainty avoidance with values over 80: Malta, Ukraine, Russia, El Salvador, Japan, Serbia, Romania, Armenia, Slovenia, Bosnia and Herzegovina, Turkey, Georgia, Korea, Hungary and Israel. On the other hand, countries with low uncertainty avoidance index are more relaxed and easily tolerate deviation from norms. They appreciate innovation and believe that rules which do not work should be changed. For instance, low uncertainty avoidance countries (0 up to 50) are Singapore, Hong Kong, England, Scotland, Malaysia, the US, Indonesia and Lebanon.

5.3.2 Across-country differences: traditional teaching methods

In this work, along with modern teaching methods we also investigate the differences in the role of traditional teaching methods on students' test scores across countries. In order to identify what country characteristics may explain the differences in the role of traditional teaching practices, we performed a separate BMA analysis. Unlike in the case of modern teaching methods, the results are less straightforward, and the best models do not provide a truly strong evidence which variables to include along with GDP per capita, measured in PPP. For more details, see Chapter 5.4.4.

Table 9: Results from second-stage international analysis explaining international differences in the estimated impact of traditional teaching practices on student performance

	Estimated effect of traditional teaching index							
Independent variable	**(1)**		**(2)**		**(3)**		**(4)**	
Purchasing power parity (GDP per capita, PPP)	-0.00001	***	-0.00001	***	-0.00001	***	-0.00001	***
	0.00000		*0.00000*		*0.00000*		*0.00000*	
Uncertainty avoidance index					-0.010	***	-0.012	***
					0.002		*0.003*	
Control of corruption	0.296	***						
	0.073							
Advanced economies			0.468	***				
			0.125					
Life expectancy (years)					0.048	***		
					0.012			
No. of observations	43		43		38		38	
R squared	0.439		0.414		0.666		0.496	

Note: Standard errors (SE) are reported in italics. Standard errors of the estimated effect of the traditional teaching practices on student performance from the first-stage country-level analysis are used as weights. * p < 0.1; ** p < 0.05; *** p < 0.01

Estimation results investigating the differences in the effect of *traditional* teaching methods are presented in Table 9. As the evidence is not entirely straight, we report the results from best 3 models and a model including variables *GDP_PPP* and *UNA,* as the latter variable is suggested to be one of the most important explanatory variables by BMA. Furthermore, cultural dimension uncertainty avoidance (*UNA*) is preferred when taking into consideration correlation between variables, as argued in Chapter 5.4.4. The interpretation of the coefficients is the same as in case of the second-stage international analysis explaining the differences in the effect of modern teaching methods in previous session.

Our results suggest that GDP per capita, measured in purchasing power parity (PPP) is negatively correlated with the effect of the traditional teaching practices on students' performance. However, we take this variable more as a control variable, as the effect seems to be driven by rich Middle East countries, namely Qatar and Kuwait.

In addition, the findings suggest that uncertainty avoidance as a cultural dimension is negatively correlated with the estimated effect of traditional teaching methods on students' test scores. As explained, this finding is intuitive. In countries with low uncertainty avoidance index, i.e. in countries which are more open-minded and do not require people to follow a rigid code of behaviour, traditional teaching methods such as lecturing can be perceived as a more unique method and easily draw students' attention and in turn, positively influence students' test scores. In other words, in more rigid countries there is a lower effect of traditional teaching methods on student performance.

In case of control of corruption (*corrup*), advanced economies (*advanced*) classification and life expectancy at birth (*life*) we find a positive correlation with the effect of traditional teaching practices on student's test scores. In fact, these variables are strongly correlated between each other, see Appendix, and thus their impact is very similar. Furthermore, the first two variables are negatively correlated with *UNA* and thus we prefer *UNA* out of these three variables. Looking at the positive estimated effect of advanced economies, our finding is in line with Bietenbeck's study (2014) who suggests a positive effect of traditional teaching methods on students' performance in a pooled sample of 9 advanced economies. However, as explained, he performed this analysis as a robustness check and did not look at the countries individually.

5.3.3 Policy recommendations

In conclusion, we have identified uncertainty avoidance as a key country characteristic which assists in explaining the international differences in the effect of both modern and traditional teaching practices on student performance in standardized tests.

We have identified that uncertainty avoidance is positively related to the effect of modern teaching methods on students' test scores. Translating this into policy recommendations, our results would suggest that policy makers in countries with higher uncertainty avoidance index should rather promote modern teaching methods in order to improve student performance. These countries prefer planning ahead and rigid codes of behaviour: e.g. Qatar, Kuwait, Russia, Bahrain, Japan, Jordan, Serbia or Ukraine.

On the other hand, our findings indicate that uncertainty avoidance is negatively related to the effect of traditional teaching methods on students' test scores. Policy makers in countries with low uncertainty index, i.e. more flexible countries who easily tolerate deviation from norms, will more likely benefit from the promotion of traditional teaching methods in order to improve student achievement in standardized tests. The latter policy recommendation applies, for instance, in England, Hong Kong, Malaysia, Scotland or the United States.

5.4 Robustness checks

5.4.1 First-stage analysis: robustness to alternative definitions of teaching practices

In the first-stage analyses, we perform a robustness check that our estimates from country level analyses are robust to alternative measures of teaching practices and not sensitive to a particular choice of measurement. We follow Bietenbeck (2014) and Korbel & Paulus (2018).

First, we define an alternative specification of the teaching index. A modern practice to solve complex problems in math and a traditional teaching method reading textbooks in science is added, respectively. In vast majority of countries (77%), our results are robust, and the measured effects exhibit the same sign and statistical significance level. These results are available per request.

Second, we address possible misspecification of teaching practice work problems on your own as one can consider it as either traditional or modern teaching method in a different situation. Thus, we replace this practice with writing equations and functions in math and reading textbooks in science classes. Again, we obtain alike results in most countries from our sample (74%). Results are available per request. We call for future research to clarify the exact effect and role of individual dimensions of teaching index and in particular contexts.

Last but not least, we included other teaching methods in our first-stage country-level analysis, such as reviewing homework or having test. The focus of this work is to examine the effects of modern and traditional teaching methods on test scores and then compare the results internationally. We observe a positive correlation between the current level of the use of modern and traditional teaching practices in a classroom on a country level which inquires how our results should be interpreted. In our country-level preferred model specification, we include both traditional and modern teaching methods index in the same regression. Our coefficients of interest are interpreted as the effect of the increase in modern / traditional teaching methods on student performance, holding traditional / modern teaching methods constant. In this scenario, we do not consider increase in other classroom practices such as reviewing homework and having a quiz or increase in modern and traditional teaching methods at the cost of the other practices. Thus, in practice, adding other teaching methods in the regression along with traditional and modern teaching methods would mean that the total teacher productivity needs to increase. Such a model specification serves as an additional robustness check. This analysis was performed only for countries where data is available. Results from these unreported regressions were similar to our main model specification.

5.4.2 First-stage analysis: complementarities between modern and traditional methods

As discussed above, we argue that modern and traditional teaching practices are complementary. This is supported by the fact that positive correlation between modern and traditional teaching methods used in a classroom was observed or all countries in our sample. However, as some previous studies suggest these teaching methods to be complementary (Klaveren 2011; Schwerdt & Wuppermann 2011), we have estimated the regression for solely modern teaching methods and solely traditional teaching methods. In these single-treatment models on a country level, a drop in the size of the coefficients was observed when both methods are included in vast majority of countries, meaning that both teaching methods are inter-related, and each method captures a part of the effect of the other method. Thus, our preferred model specification includes both teaching methods.

5.4.3 Second-stage international analysis: Bayesian model averaging (BMA)

As explained, it is not known in advance what country characteristics should be included in our second-stage model explaining the international differences in the effect of modern and traditional teaching practices on students' test scores, because no specific literature exists. Thus, we perform Bayesian model averaging (BMA) method, which assists us with a variable selection and a particular model choice. BMA is followed separately for modern and separately for the effect of traditional teaching practices. The method tackles the challenges we are facing: (1) limited evidence about what variables should be included in the true model and (2) limited number of observations. BMA method provides weighted average of all possible combinations of models, justifying a particular model choice (Zeugner, 2011; Banner & Higgs 2017).

We choose a default option of model averaging using BMS package in R to examine the ability of 28 potential variables to predict the differences in the effect of modern and traditional teaching methods on student test scores across countries (Hoeting et al. 1999; Zeugner, 2011). We divide our dependent variable by the estimated standard error from the first-stage country-level analysis, as Bayesian model averaging does not allow including weights.

First, looking at the international differences in the modern teaching methods, the results of BMA are presented below and suggest two variables to be unambiguously included in the true model: cultural dimension uncertainty avoidance (*UNA*) and gross domestic product per capita, measured in PPP (*GDP_PPP*) with posterior inclusion probabilities (PIP) of 95.3% and 82.6% that a given variable belongs to the final model, respectively (Hoeting et al. 1999).

Furthermore, the BMA suggests mean educational achievement in years in a given country (*mean_y_school*) to be included in the true model with a posterior inclusion probability 33.9%. See Table 11 and Figure 2. Additionally, the results show that the model size should be around 3.5 variables, see Figure 1. Lastly, the method offers graphical analysis suggesting a particular model inclusion, more details can be found in Figure 2.

The best model includes cultural dimension uncertainty avoidance (*UNA*) and gross domestic product per capita, measured in PPP (*GDP_PPP*) with posterior model probability (PMP) of 18.3% indicating what proportion of total PMP the top model accounts for. The second-best model indicates the two aforementioned variables along with mean educational achievement in years in a given country (*mean_y_school*) with a similar PMP of 16.5%, as shown in Table 12. We present these two models with these three variables as our preferred model specifications of the weighted least squares (WLS) model in this work.

Furthermore, these three variables are included in majority model mass. The BMA results also suggest a sign of the coefficients, i.e. sign certainty. The value one in the last column in Table 11 indicates a positive sign of the estimated coefficient, value zero suggests a negative sign. We can see that *UNA* and *GDP_PPP* is suggested to have a positive sign in the true model in all cases while *mean_y_school* estimates indicate to have a negative sign.

In regard to other variables, BMA method indicates that the proportion of migrant population of total population (*migrant_stock*) and current level of modern teaching practices used in a classroom (*index_mod*) to be included in the true model with PIP over 15%, see Table 12. However, it is important to highlight that these additional variables are usually not in the model at the same time. For instance, the third-best model includes *UNA* and *migrant_stock* replacing *GDP_PPP*. These latter two variables are strongly correlated of 0.82 and thus, it is preferrable to work with *GDP_PPP* which exhibits higher values of PIP than *migrant_stock*. Similar reasoning applies in case of the *index_mod* variable appearing in the fourth-best model which includes UNA, GDP_PPP and *index_mod* instead of *mean_y_school*. Again, we observe a relatively strong correlation of -0.431 and therefore, we prefer to work the variable *mean_y_school* instead of *index_mod* which exhibits a higher level of PIP.

To conclude, BMA deals with model uncertainty issue and supports us with a particular variable selection and a preferred model choice in the second-stage analysis. However, due to complexity of interpretation of model averaging method, we decided to present results of WLS model in this work.

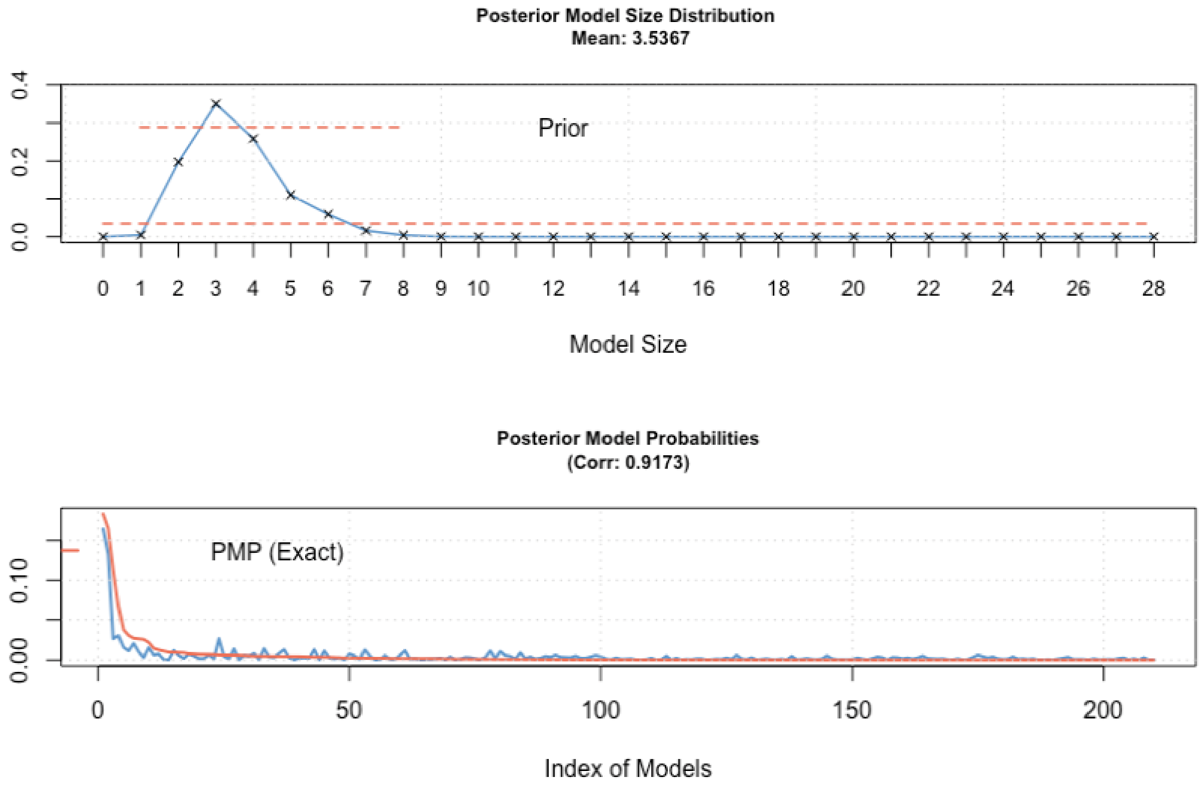

Figure 1: BMA: model size and convergence

Table 10: BMA: model diagnostics

Mean no. regressors	**Draws**	**Burnins**	**Time**
3.50	3,000	1,000	0.32 secs
No. models visited	**Model space 2^K**	**% visited**	**% topmodels**
507	2.70E+08	0.00019	100
Corr PMP	**No. obs.**	**Model prior**	**g-prior**
0.927	36	random / 14	UIP
Shrinkage-stats			
Av=0.973			

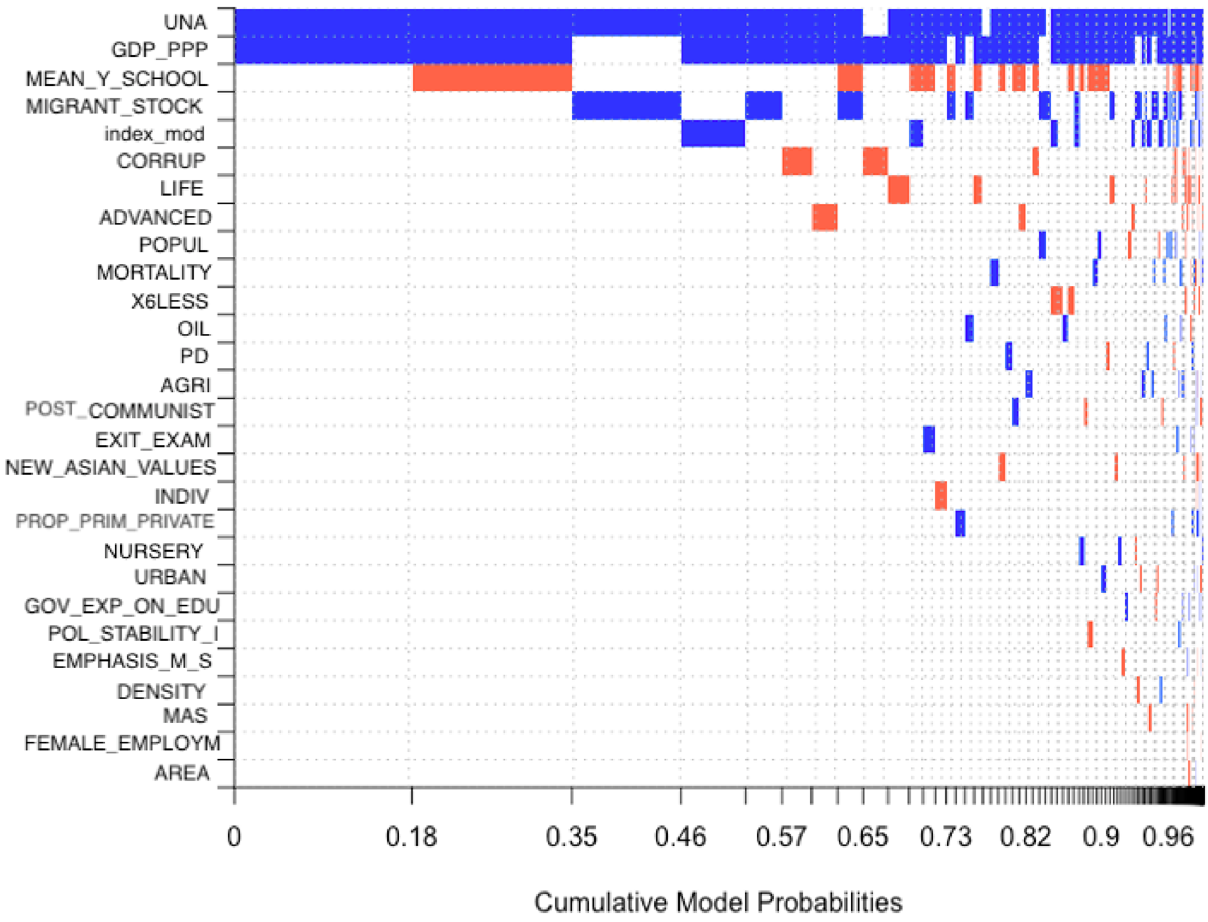

Figure 2: BMA: model inclusion based on best 224 models

Table 11: Marginal posterior summaries of coefficient under BMA

Independent variable	PIP	Post mean	Post SD	Cond. Pos. Sign
UNA	0.953	0.043	0.015	1.000
GDP_PPP	0.826	0.000	0.000	1.000
MEAN_Y_SCHOOL	0.339	-0.085	0.134	0.000
MIGRANT_STOCK	0.289	0.020	0.034	1.000
index_mod	0.173	1.088	2.727	1.000
LIFE	0.091	-0.008	0.032	0.000
MORTALITY	0.076	0.000	0.006	0.719
CORRUP	0.066	-0.047	0.213	0.000
POST_COMMUNIST	0.059	-0.003	0.181	0.225
PD	0.058	0.000	0.004	0.684
ADVANCED	0.058	-0.049	0.268	0.000
POPUL	0.056	0.000	0.001	0.763
ARGI	0.055	0.001	0.005	0.933
NEW_ASIAN_VALUES	0.051	-0.021	0.166	0.000
GOV_EXP_ON_EDU	0.048	0.004	0.048	0.720
NURSERY	0.044	0.002	0.131	0.450
DENSITY	0.040	0.000	0.000	0.550
OIL	0.039	0.001	0.008	0.983

EXIT_EXAM	0.034	0.027	0.198	1.000
EMPHASIS_M_S	0.029	-0.001	0.089	0.364
INDIV	0.027	0.000	0.003	0.400
PROP_PRIM_PRIVATE	0.024	0.000	0.003	1.000
X6LESS	0.017	-0.015	0.153	0.000
URBAN	0.016	0.000	0.003	0.367
MAS	0.012	0.000	0.002	0.000
FEMALE_EMPLOYMENT	0.008	0.000	0.003	0.167
POL_STABILITY_I	0.008	0.000	0.029	0.652
AREA	0.002	0.000	0.000	0.800

Table 12: Overview of inclusion of TOP 10 models under BMA

Variable	TOP 1	TOP 2	TOP 3	TOP 4	TOP 5	TOP 6	TOP 7	TOP 8	TOP 9	TOP 10
index_mod	0	0	0	1	0	0	0	0	0	0
EXIT_EXAM	0	0	0	0	0	0	0	0	0	0
PD	0	0	0	0	0	0	0	0	0	0
INDIV	0	0	0	0	0	0	0	0	0	0
MAS	0	0	0	0	0	0	0	0	0	0
UNA	1	1	1	1	1	1	1	1	0	1
NEW_ASIAN_VALUES	0	0	0	0	0	0	0	0	0	0
POPUL	0	0	0	0	0	0	0	0	0	0
AREA	0	0	0	0	0	0	0	0	0	0
DENSITY	0	0	0	0	0	0	0	0	0	0
URBAN	0	0	0	0	0	0	0	0	0	0
LIFE	0	0	0	0	0	0	0	0	0	1
MORTALITY	0	0	0	0	0	0	0	0	0	0
GDP_PPP	1	1	0	1	1	1	1	1	1	1
GOV_EXP_ON_EDU	0	0	0	0	0	0	0	0	0	0
EMPHASIS_M_S	0	0	0	0	0	0	0	0	0	0
NURSERY	0	0	0	0	0	0	0	0	0	0
X6LESS	0	0	0	0	0	0	0	0	0	0
POST_COMMUNIST	0	0	0	0	0	0	0	0	0	0
POL_STABILITY_I	0	0	0	0	0	0	0	0	0	0
CORRUP	0	0	0	0	0	1	0	0	1	0
ARGI	0	0	0	0	0	0	0	0	0	0
FEMALE_EMPLOYMENT	0	0	0	0	0	0	0	0	0	0
MIGRANT_STOCK	0	0	1	0	1	0	0	1	0	0
OIL	0	0	0	0	0	0	0	0	0	0
PROP_PRIM_PRIVATE	0	0	0	0	0	0	0	0	0	0
MEAN_Y_SCHOOL	0	1	0	0	0	0	0	1	0	0
ADVANCED	0	0	0	0	0	0	1	0	0	0
PMP (Exact)	0.183	0.165	0.112	0.067	0.038	0.030	0.027	0.026	0.026	0.022
PMP (MCMC)	0.085	0.101	0.064	0.048	0.029	0.016	0.001	0.009	0.003	0.038

5.4.4 Second-stage international analysis: Bayesian model averaging (BMA) for traditional teaching practices

In the previous part. we have performed BMA method in order to find out the most important explanatory variables which should be included in the true model explaining the differences across countries in the role of *modern* teaching practices on students' test scores. However. the differences across countries in the role of *traditional* teaching methods might be explained by different country characteristics which should be included in the true model. Therefore, the BMA analysis was performed separately for traditional teaching methods and the outcomes are reported below, see Figures 3 and 4 and Tables 13-15.

The findings are less straightforward than in the previous case. BMA method clearly suggests only one particular variable to be included in the true model: gross domestic product per capita, measured in PPP (*GDP_PPP*) with posterior inclusion probability (PIP) of 95.1%, see Table 14. Furthermore, BMA indicates uncertainty avoidance (*UNA*) with PIP of 41.6% to be included in the true model. In addition, it suggests control of corruption (*corrup*), advanced economies classification (*advanced*), infant mortality rate as per 1,000 live births (*mortality*) and life expectancy at birth (*life*) to be included in the true model with PIP over 15%.

However, these variables are usually not included at the same time in the best models. Our intuition advises these variables are correlated. In fact, we observe rather strong (negative) correlations: 0.78 between *corrup* and *advanced*, -0.53 between *corrup* and *UNA*, and -0.42 between *advanced* and *UNA*, see Appendix. Thus, we follow similar reasoning as in the previous case and prefer *UNA* variable to be included in our true model out of these three, as its PIP exhibits highest value of 0.416, see Table 14.

In case of mortality, the BMA method provides unclear suggestions as PIP value is quite high of 23.7% but the variable is included only in models with posterior model probability (PMP) less than 3%. Similarly, BMA method does not strongly support variable *life* to be included in the true model. PIP of the variable is over 15% but PMP of the best models in which it appears is only around 5%.

To conclude, BMA method unambiguously suggests only variable *GDP_PPP* to be included in the true model. Then, we can argue that *UNA* is the next important variable. However, support for any additional variables is limited.

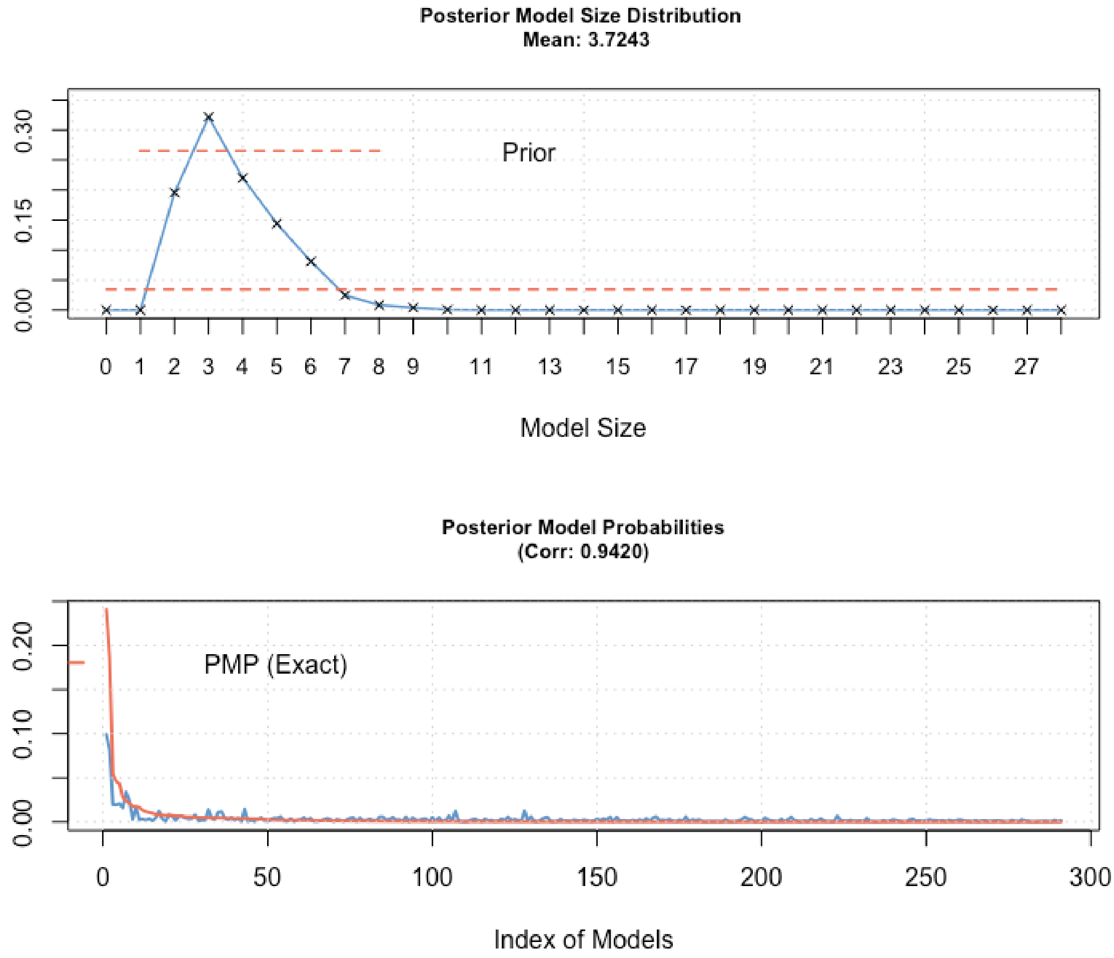

Figure 3: BMA: model size and convergence (international differences in the estimated effect of traditional teaching practices on student test scores)

Table 13: BMA: model diagnostics (international differences in the estimated effect of traditional teaching practices on student test scores)

Mean no. regressors	**Draws**	**Burnins**	**Time**
3.72	3,000	1,000	0.31 secs
No. models visited	**Model space 2^K**	**% visited**	**% topmodels**
698	2.70E+08	0.00026	100
Corr PMP	**No. obs.**	**Model prior**	**g-prior**
0.942	36	random / 14	UIP
Shrinkage-stats			
Av=0.973			

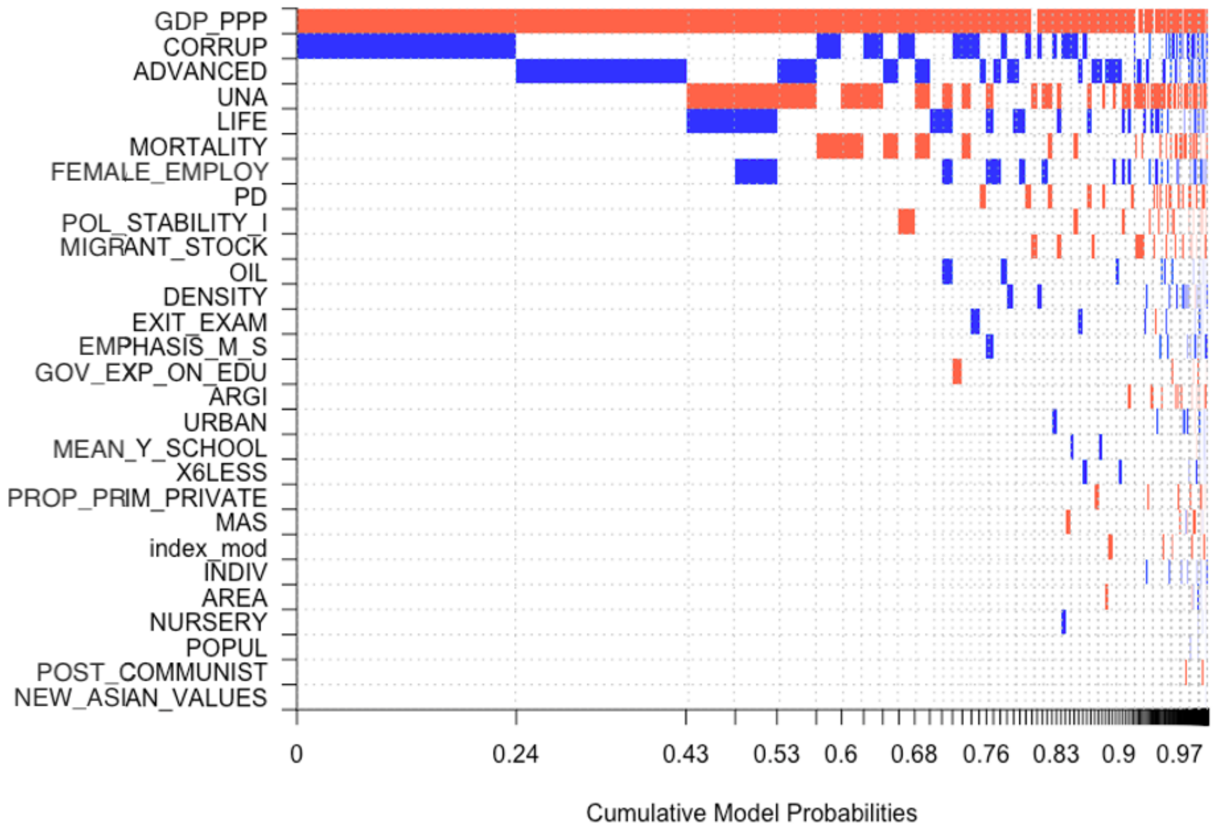

Figure 4: BMA: model inclusion based on best 291 models (international differences in the estimated effect of traditional teaching practices on student test scores)

Table 14: Marginal posterior summaries of coefficient under BMA (international differences in the estimated effect of traditional teaching practices on student test scores)

Independent variable	PIP	Post mean	Post SD	Cond. Pos. Sign
GDP_PPP	0.951	0.000	0.000	0.000
UNA	0.416	-0.019	0.026	0.000
CORRUP	0.379	0.758	1.057	1.000
ADVANCED	0.311	0.902	1.487	1.000
MORTALITY	0.237	-0.015	0.031	0.000
LIFE	0.188	0.048	0.112	1.000
FEMALE_EMPLOYMENT	0.147	0.012	0.032	1.000
PD	0.134	-0.004	0.013	0.000
MIGRANT_STOCK	0.088	-0.007	0.025	0.000
POL_STABILITY_I	0.084	-0.060	0.264	0.000
ARGI	0.081	-0.002	0.013	0.062
DENSITY	0.075	0.000	0.000	0.991
URBAN	0.071	0.002	0.011	0.925
OIL	0.067	0.003	0.016	1.000
EXIT_EXAM	0.062	0.043	0.297	0.903
index_mod	0.057	-0.120	1.445	0.082
X6LESS	0.050	0.040	0.314	1.000
EMPHASIS_M_S	0.050	0.061	0.304	1.000
INDIV	0.046	0.001	0.007	1.000
AREA	0.040	0.000	0.000	0.600

PROP_PRIM_PRIVATE	0.038	-0.001	0.008	0.000
MEAN_Y_SCHOOL	0.037	0.001	0.038	0.757
MAS	0.036	0.000	0.005	0.130
GOV_EXP_ON_EDU	0.035	-0.012	0.082	0.000
POPUL	0.020	0.000	0.001	1.000
NURSERY	0.016	0.007	0.132	1.000
POST_COMMUNIST	0.007	0.000	0.076	0.318
NEW_ASIAN_VALUES	0.002	0.003	0.075	1.000

Table 15: Overview of inclusion of TOP 10 models under BMA (international differences in the estimated effect of traditional teaching practices on student test scores)

Variable	TOP 1	TOP 2	TOP 3	TOP 4	TOP 5	TOP 6	TOP 7	TOP 8	TOP 9	TOP 10
index_mod	0	0	0	0	0	0	0	0	0	0
EXIT_EXAM	0	0	0	0	0	0	0	0	0	0
PD	0	0	0	0	0	0	0	0	0	0
INDIV	0	0	0	0	0	0	0	0	0	0
MAS	0	0	0	0	0	0	0	0	0	0
UNA	0	0	1	1	1	0	1	1	0	0
NEW_ASIAN_VALUES	0	0	0	0	0	0	0	0	0	0
POPUL	0	0	0	0	0	0	0	0	0	0
AREA	0	0	0	0	0	0	0	0	0	0
DENSITY	0	0	0	0	0	0	0	0	0	0
URBAN	0	0	0	0	0	0	0	0	0	0
LIFE	0	0	1	1	0	0	0	0	0	0
MORTALITY	0	0	0	0	0	1	1	0	1	0
GDP_PPP	1	1	1	1	1	1	1	1	1	1
GOV_EXP_ON_EDU	0	0	0	0	0	0	0	0	0	0
EMPHASIS_M_S	0	0	0	0	0	0	0	0	0	0
NURSERY	0	0	0	0	0	0	0	0	0	0
X6LESS	0	0	0	0	0	0	0	0	0	0
POST_COMMUNIST	0	0	0	0	0	0	0	0	0	0
POL_STABILITY_I	0	0	0	0	0	0	0	0	0	1
CORRUP	1	0	0	0	0	1	0	1	0	1
ARGI	0	0	0	0	0	0	0	0	0	0
FEMALE_EMPLOYMENT	0	0	0	1	0	0	0	0	0	0
MIGRANT_STOCK	0	0	0	0	0	0	0	0	0	0
OIL	0	0	0	0	0	0	0	0	0	0
PROP_PRIM_PRIVATE	0	0	0	0	0	0	0	0	0	0
MEAN_Y_SCHOOL	0	0	0	0	0	0	0	0	0	0
ADVANCED	0	1	0	0	1	0	0	0	1	0
PMP (Exact)	0.241	0.187	0.054	0.046	0.043	0.028	0.024	0.021	0.018	0.017
PMP (MCMC)	0.099	0.081	0.020	0.019	0.021	0.016	0.034	0.024	0.003	0.016

6 Conclusion

There is an ongoing debate about what teaching practices are the most effective ones in order to improve student learning and performance. However, little is known about the impact across countries and the evidence is highly inconclusive: some studies suggest no effect of both modern and traditional teaching practices on students' test scores (Klaveren 2011; Korbel & Paulus 2018), some studies suggest a positive effect of traditional teaching practices (Schwerdt & Wuppermann 2011; Bietenbeck 2014) and lastly, some studies report a positive effect of modern teaching practices on students' test scores (Aslam & Kingdon 2011; Lavy 2016). Thus, we can expect significant differences in the role of modern (or traditional) teaching methods on student performance across countries (Hypothesis 1). We extend the analysis to 43 countries for which the TIMSS data is available to provide more consistent international evidence and examine why these differences across countries exist (Hypothesis 2).

Our analysis is performed in two steps and is a typical example of hierarchical linear modelling (HLM). In the first stage, we investigate the effect of modern and traditional teaching practices on student test scores on a country level. We use student fixed effect strategy in order to control for majority of selection issues, e.g. as students are not randomly assigned to classes and schools.

Our country-level results provide evidence suggesting there is indeed no one-fits-to-all approach towards the use of modern and traditional teaching methods to improve student performance across countries. First, we find positive effect of modern teaching practices on student performance in 10 countries: Bahrain, Bosnia and Herzegovina, Japan, Jordan, Kuwait, Qatar, Russian Federation, Serbia, Tunisia and Ukraine. Second, we found negative effect of modern teaching methods on students' test scores in Malaysia.

On the other hand, our results suggest a positive effect of traditional teaching methods on student performance in 12 countries: Australia, Taiwan, England, Hong Kong, Israel, Korea, Lebanon, Lithuania, Malaysia, Scotland, Thailand and the US. In contrast, we identified a negative effect of traditional teaching methods on student performance in Qatar.

Furthermore, we have identified both modern and traditional teaching methods to have a positive effect on student performance in two countries, but the effect is stronger for traditional teaching methods: Colombia and Singapore. One possible explanation is that in these countries other teaching practices such as having test or quiz, using computers, doing and reviewing homework, practicing calculations, reading textbook etc. are most prevalent. These methods are not classified as either modern or traditional teaching method by our definition.

Lastly, in almost half of the countries (19) in our sample, we found neither teaching practices have any effect on student performance. This is the case of Armenia, Botswana, Cyprus, Czech Republic, Egypt, El Salvador, Georgia, Ghana, Hungary, Indonesia, Iran, Malta, Mongolia, Oman, Palestinian Nat'l Auth., Romania, Saudi Arabia, Slovenia and Turkey.

As we identified differences across countries, we can continue our investigation and try to identify factors explaining these differences across countries. Following previous literature and economic intuition, author collected unique data set containing 28 country characteristics which could potentially explain the international differences. Bayesian model averaging (BMA) method assisted in dealing with model uncertainty and supported our preferred model choice. The second stage preferred model specification was estimated using weighted least squares (WLS) method which allows more straightforward interpretation of our results. We use quality of the estimated effect from the first-stage country level analysis as weights.

Our most important finding is that cultural dimension uncertainty avoidance assists in explaining the international differences in the effect of modern and traditional teaching practices on student performance in standardized tests. Our results suggest that uncertainty avoidance index is positively correlated with the effect of modern teaching practices on student performance. This may be interpreted that policy makers in countries with higher uncertainty avoidance index should rather promote modern teaching methods in order to improve student performance. Uncertainty avoidance countries rather prefer planning and rigid codes of behaviour. These are, for example, Qatar, Kuwait, Russia, Bahrain, Japan, Jordan, Serbia or Ukraine.

Looking at *traditional* teaching methods and the observed international differences, our findings suggest that uncertainty avoidance is negatively related to the effect of traditional teaching methods on students' test scores. Translating this into policy recommendations, policy makers in countries with low uncertainty index, i.e. more flexible and open-minded countries who easily tolerate deviation from norms, will more likely benefit from the promotion of traditional teaching methods in order to improve student achievement in standardized tests. This policy recommendation applies to, for instance, England, Scotland, the United States, Hong Kong and Malaysia.

However, our work has also some limitations which should be taken into account when interpreting our results. Our findings are based on the analysis of math and science and in our sample, we have a lot of non-advanced economies and countries from the Middle East. The relationship may be different for different subjects and a different sample of countries. Besides, although student fixed effects approach accounts for majority of potential bias due to selection issues, it is based on an assumption that students do not sort into teaching practices in a subject specific way. This would be violated if students are sorted into classes and teaching practices focusing on a particular subject. However, we focus on eight-grade students and student tracking usually takes place later on (Hanushek & Woessmann 2010). Furthermore, we focus on the impact of modern and traditional teaching methods test scores in this work, but research suggest that modern teaching methods can develop skills which are not captured by test scores, such as cognitive skills or a skill of teamwork (Algan et al. 2013; Bietenbeck 2014). Future research should investigate a role of these skills.

Despite the limitations, our study contributes to existing literature in several ways. First, we contribute to current debate about the role of modern and traditional teaching methods on student performance. Our work allows for immediate comparison of the effect of modern and traditional teaching practices on student performance across countries, as the same methodology, data and definition of teaching methods are used. Second, this works provides international evidence for 43 countries. To author's best knowledge, this is the first paper that examined why these differences across countries exist. Our findings can be of a use to policy makers and teachers in order to better understand the role of a particular teaching method in order to improve student performance in standardized tests. Lastly, we include culture in order to explain the across country differences which is a first study of its kind (Hanushek & Woessmann 2010).

References

Abadie, A., S. Athey, G. W. Imbens & J. Wooldridge (2017): “When should you adjust standard errors for clustering?” *NBER Working Papers 24003*, National Bureau of Economic Research. https://doi.org/10.3386/w24003

Aslam, M., & G. Kingdon (2011): “What Can Teachers Do to Raise Pupil achievement?” *Economics of Education Review* **30(3):** pp. 559–574. doi:10.1016/j.econedurev. 2011.01.001

Banner, K. M. & M. D. Higgs (2017): “Considerations for assessing model averaging of regression coefficients.” *Ecological Applications* **27(1):** pp. 78-93

Bietenbeck, J. (2014): “Teaching Practices and Cognitive Skills.” *Labour Economics* **30:** pp. 143-153. doi:10.1016/j.labeco.2014.03.002

Bishop, J. H. (1997): “The effect of national standards and curriculum-based exams on achievement.” *The American Economic Review* **87(2):** pp. 260-264

Bishop, J. H. (2005): “High School Exit Examinations: When Do Learning Effects Generalize?” *CAHRS Working Paper 05-04*, Cornell University

Cascio, E., C. Damon & N. Gordon (2008): "Education and the age profile of literacy into adulthood." *Journal of Economic Perspectives* **22(3):** pp. 47-70

Chang, H. (2002): *Kicking Away the Ladder: Development Strategy in Historical Perspective.* London Anthem Press

CPDS (2008): Comparative Data Set for 28 Post-Communist Countries. Available at: https://cpds-data.org/images/Archive/CPDSII/ExcelCodebook2008_ger.pdf

Donald, S. G. & K. Lang. (2007): “Inference with difference-in-differences and other panel data.” *The Review of Economics and Statistics* **89(2):** pp. 221-233

Foy, P. & J. F. Olson (2009): “TIMSS 2007 User Guide for the International Database.” *TIMSS 2007,* Chestnut Hill, TIMSS & PIRLS international Study Center, Boston College. ISBN: 1-889938-51-3

Global Economy (2020): *Business and economic data for 200 countries.* Available at: https://www.theglobaleconomy.com

Hamilton, L. S., D. F. McCaffrey, B. M. Stecher, S. P. Klein, A. Robyn & D. Bugliari (2003): "Studying large-scale reforms of instructional practice: an example from mathematics and science." *Educational Evaluation and Policy Analysis* **25(1):** pp. 1-29

Hanushek, E. A. & L. Woessmann (2011): "The economics of international differences in educational achievement." *Handbook of the Economics of Education* **3:** pp. 89-200

Harmouch, T., Z. Khraibani & T. Atrissi (2017): "A Multilevel Analysis to Analyse the TIMSS Data: A Comparison of The Lebanese And Singapore." *International Journal of New Technology and Research* **3(9):** pp. 87-102

Hoeting, J. A., D. Madigan, A. E. Raftery & C. T. Volinsky (1999): "Bayesian model averaging: a tutorial." *Statistical Science* **14(4):** pp. 382-401

Hofmann, D. A. (1997): "An overview of the logic and rationale of hierarchical linear models." *Journal of Management* **23(6):** pp. 723-744

Hofstede, G. (1991): *Cultures and organizations: Software of the mind*. London McGraw-Hill

Hofstede, G. (2020): *Hofstede Insights.* Available at: https://www.hofstede-insights.com/country-comparison/

Hoxby, C. M. H. U. (2000): "Peer Effects in the Classroom: Learning from Gender and Race Variation." *NBER Working Papers 7867*. National Bureau of Economic Research. https://doi.org/10.3386/w7867

Huang, M. (2009): "Do public exams raise student performance? A cross-national difference-in-differences analysis." *Journal of Social Sciences and Philosophy* **21(1):** pp. 1-24

IMF (2007): *World Economic Outlook, October 2007.* IBSN 978-1-58906-688-5

Joncas, M. (2008): "TIMSS 2007 Sample Design." In Olson, J. F., M. O. Martin & I. V. S. Mullis (Eds.): TIMSS *2007 Technical Report (pp. 77-92).* Chestnut Hill, TIMSS & PIRLS international Study Center, Boston College

Klaveren, C. V. (2011): "Lecturing Style Teaching and Student Performance." *Economics of Education Review* **30(4):** pp. 729-739. doi:10.1016/j.econedurev.2010.08.007

Koch, A., J. Nafziger & H. S. Nielsen (2015): "Behavioral Economics of Education." *Journal of Economic Behavior & Organization* **115:** pp. 3-17. https://doi.org/10.1016/j.jebo.2014.09.005

Korbel, V. & M. Paulus (2018): “Do teaching practices impact socio-emotional skills?“ *Education Economics* **26(4):** pp. 337-355. https://doi.org/10.1080/09645292.2018.1460320

Krueger, A. B. (1999): “Experimental estimates of education production functions.” *The quarterly journal of economics* **114(2):** pp. 497-532

Kunter, M. & J. Baumert (2006): “Who is the expert? Construct and criteria validity of student and teacher ratings of instruction.” *Learning Environments Research* **9(3):** pp. 231-251

Lavy, V. (2016): “What Makes an Effective Teacher? Quasi-experimental Evidence.” *CESifo Economic Studies* **62(1):** pp. 88-125. https://doi.org/10.1093/cesifo/ifv001

Martin M.O., I. V. Mullis & P. Foy (2008). “TIMSS 2007 International mathematics report. Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.” *TIMSS 2007*, Chestnut Hill, TIMSS & PIRLS international Study Center, Boston College. Available at: https://timssandpirls.bc.edu/TIMSS2007/context.html

Mensah, Y. M. & H. Y. Chen (2013): “Global clustering of countries by culture: an extension of the GLOBE study.” *SSRN*. Available at: http://dx.doi.org/10.2139/ssrn.2189904

Mons, N. (2009): *Theoretical and real effects of standardised assessment.* Eurydice

Montgomery, J. L. & W. Baker (2007): “Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance.” *Journal of Second Language Writing* **16(2):** pp. 82-99

Mullis, I. V. S, M. O. Martin, J. F. Olson, D. R. Berger, D. Milne & G. M. Stanco (2008): “TIMSS 2007 Encyclopedia: A Guide to Mathematics and Science Education Around the World, Volumes 1 and 2”. *TIMSS 2007*, Chestnut Hill, TIMSS & PIRLS international Study Center, Boston College. Available at: 1. https://timssandpirls.bc.edu/TIMSS2007/encyclopedia.html

National Research Council (1996): *National Science Education Standards.* The National Academies Press, Washington, DC. https://doi.org/10.17226/4962

Resnick, L. (1987): *“Education and Learning to Think.*” National Academy Press, Washington, DC

Rivkin, S. G., E. A. Hanushek & J. F. Kain (2005): “Teachers, schools, and academic achievement.” *Econometrica* **73(2):** pp. 417-458

Schuetz, G. (2009): "Does the quality of pre-primary education pay off in secondary school? An international comparison using PISA 2003." *IFO Working Paper 68*. IFO Institute for Economic Research

Schwerdt, G., & A. C. Wuppermann (2011): "Is Traditional Teaching Really all that Bad? A Within-student Between-subject Approach." *Economics of Education Review* **30(2):** pp. 365–379. doi:10.1016/j.econedurev.2010.11.005

United Nations database (2020), "UIS Statistics". *UNESCO*. Available at: http://data.uis .unesco.org

Varkey Foundation (2018): *The Global Teacher Status Index*. Available at: https://www.varkeyfoundation.org/media/4790/gts-index-9-11-2018.pdf

Walker, P. (2012): "Tough exams and learning by rote are the keys to success, says Michael Gove." *The Guardian*

Wenglinsky, H. 2002. "How schools matter: The link between teacher classroom practices and student academic performance." *Education Policy Analysis Archives* **10(12).** doi:10.14507/epaa.v10n12.2002

Wigfield, A., K. R. Wentzel & D. Miele (2009): *Handbook of Motivation at School*. Routledge, NY.

Woessmann, L. (2016): "The importance of school systems: Evidence from international differences in student achievement." *Journal of Economic Perspectives* **30(3):** pp. 3-32

Woltman, H., A. Feldstain, J. C. MacKay & M. Rocchi (2012): "An Introduction to Hierarchical Linear Modeling." *Tutorials in Quantitative Methods for Psychology* **8(1):** pp 52-69

World Bank database (2020): "Education statistics." *The World Bank Group.* Available at: http://datatopics.worldbank.org/education/indicators

Zeugner, S. (2011): "Bayesian model averaging with BMS." *The Comprehensive R Archive Network.* Available at: https://cran.r-project.org/web/packages/BMS/vignettes/bms.pdf

Appendix A

In this work, we were working with student test scores which were standardized to have zero mean and standard deviation one in each country. However, it is interesting to look at the overall test scores in each country. Table A.1 below reports mean and standard deviation of the test scores in each country. As suggested in the User Guide by TIMSS (Foy & Olson 2009), the mean test score in each country was calculated from the 5 plausible values, using students' weights.

Table A.1: Test scores in each country

	Test scores													
	Overall		Math		Science		Biology		Chemistry		Earth science		Physics	
	Mean	**SD**	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Armenia	487.47	104.77	501.14	83.82			486.60	101.39	477.81	122.53	466.75	113.05	505.06	103.08
Australia	506.81	79.63	497.27	79.21	516.34	80.05								
Bahrain	435.83	83.45	401.09	82.53	470.56	84.37								
Bosnia and Herzegovina	466.64	86.93	461.25	75.40			467.99	83.43	469.73	89.49	469.66	100.60	464.58	85.75
Botswana	362.93	87.32	367.35	75.56	358.50	99.08								
Chinese Taipei	580.26	97.55	598.58	105.41	561.93	89.69								
Colombia	401.22	76.19	382.58	77.17	419.85	75.21								
Cyprus	459.40	93.24	470.31	86.54					452.13	100.63	457.08	93.78	458.08	91.99
Czech Republic	532.77	78.06	505.43	72.97			534.75	75.34	541.65	80.61	539.91	83.86	542.10	77.52
Egypt	406.82	97.62	398.32	97.85	415.32	97.39								
El Salvador	366.63	70.27	343.64	71.37	389.62	69.17								
England	514.71	83.95	514.71	83.04	549.78	84.86								
Georgia	421.45	97.28	418.94	93.55			428.96	84.32	416.73	110.95	422.44	96.24	420.16	101.33
Ghana	310.70	98.80	313.20	90.11	308.20	107.48								
Hong Kong SAR	552.62	86.84	574.02	93.18	531.22	80.49								

Hungary	535.77	*86.66*	519.02	*84.11*			535.94	*78.05*	542.41	*91.91*	536.08	*91.88*	545.41	*87.36*
Indonesia	414.36	*85.66*	399.12	*87.40*			421.97	*84.14*					421.99	*85.45*
Islamic Rep. of Iran	432.08	*83.48*	404.39	*85.84*	459.77	*81.11*								
Israel	469.06	*97.04*	469.08	*95.73*	469.03	*98.34*								
Japan	564.69	*81.37*	571.53	*84.72*	557.85	*78.02*								
Jordan	460.33	*95.90*	432.84	*98.36*	487.81	*93.43*								
Korea	575.98	*84.19*	597.78	*91.75*	554.18	*76.63*								
Rep. of Kuwait	389.90	*82.38*	357.47	*77.31*	422.33	*87.45*								
Lebanon	435.06	*97.70*	454.42	*74.16*			408.26	*105.50*	448.07	*115.34*			429.50	*95.78*
Lithuania	515.69	*86.17*	508.48	*78.70*			532.86	*85.99*	509.20	*89.71*	519.10	*93.99*	508.79	*82.47*
Malaysia	472.56	*83.71*	474.05	*79.28*	471.06	*88.14*								
Malta	504.93	*103.69*	491.71	*92.24*			538.19	*93.73*	590.72	*95.85*	432.02	*128.97*	472.02	*107.64*
Mongolia	442.61	*86.27*	434.56	*81.25*			453.61	*81.94*	438.47	*92.47*	428.88	*98.67*	457.51	*77.01*
Oman	402.34	*92.77*	376.96	*92.89*	427.71	*92.65*								
Palestinian Nat'l Auth.	392.06	*103.33*	373.37	*99.17*	410.74	*107.48*								
Qatar	317.29	*107.98*	311.23	*91.39*	323.34	*124.56*								
Romania	464.81	*97.03*	466.93	*97.20*			461.89	*91.99*	463.08	*103.71*	471.91	*97.37*	460.26	*94.89*
Russian Federation	528.24	*85.49*	515.42	*81.22*			529.17	*80.52*	542.88	*90.30*	530.28	*86.55*	523.46	*88.88*
Saudi Arabia	370.02	*76.24*	332.60	*75.75*	407.44	*76.72*								
Scotland	497.01	*80.26*	488.22	*78.80*	505.79	*81.72*								
Serbia	474.58	*91.45*	491.25	*86.01*			478.27	*85.84*	467.38	*94.74*	466.78	*104.87*	469.23	*85.81*
Singapore	586.31	*97.61*	592.80	*91.97*	579.81	*103.24*								
Slovenia	528.81	*76.90*	503.17	*71.12*			533.60	*77.73*	548.52	*84.15*			529.95	*74.60*
Thailand	456.36	*87.18*	441.76	*91.65*	470.95	*82.70*								
Tunisia	434.14	*63.33*	422.00	*66.32*	446.27	*60.33*								
Turkey	444.64	*100.08*	433.59	*108.50*	455.68	*91.66*								
Ukraine	488.59	*86.68*	470.82	*82.98*			483.05	*81.44*	498.45	*92.24*	489.75	*90.05*	500.86	*86.68*
United States	514.93	*78.88*	508.74	*76.43*	521.12	*81.33*								

Note: Test scores are calculated from 5 plausible values in each subject. Student weights are used. Standard deviations (SD) in italics.

Appendix B

In Appendix B, we report R^2 from each country-level first-stage regression which investigate the impact of modern and traditional teaching methods on student test scores in each of 43 countries. See Table B.1.

Table B.1: R squared from first-stage country-level analyses

Country	R squared	Country	R squared
Armenia	0.766	Rep. of Kuwait	0.902
Australia	0.924	Lebanon	0.800
Bahrain	0.911	Lithuania	0.799
Bosnia and Herzegovina	0.757	Malaysia	0.912
Botswana	0.939	Malta	0.877
Chinese Taipei	0.937	Mongolia	0.778
Colombia	0.917	Oman	0.941
Cyprus	0.786	Palestinian Nat'l Auth.	0.942
Czech Republic	0.803	Qatar	0.878
Egypt	0.942	Romania	0.785
El Salvador	0.914	Russian Federation	0.804
England	0.944	Saudi Arabia	0.901
Georgia	0.686	Scotland	0.928
Ghana	0.904	Serbia	0.781
Hong Kong SAR	0.905	Singapore	0.943
Hungary	0.804	Slovenia	0.798
Indonesia	0.802	Thailand	0.932
Islamic Rep. of Iran	0.925	Tunisia	0.905
Israel	0.924	Turkey	0.949
Japan	0.928	Ukraine	0.783
Jordan	0.940	United States	0.924
Korea	0.922		

Appendix C

In our second-stage analysis, we looked at 28 country characteristics which could potentially explain differences across countries in the effect of the use of modern (traditional) teaching methods on student test scores. Descriptive statistics and correlation between each pair of the 28 explanatory variables is shown below in Table C.1 and Table C.2.

Table C.1: Second-stage analysis: descriptive statistics of all 28 explanatory variables

Independent variable	Min	Max	Mean	SD
index_mod	0.293	0.648	0.518	0.077
EXIT_EXAM	0	1	0.791	0.412
PD	13	100	68.158	20.096
INDIV	13	91	38.237	23.133
MAS	19	95	50.053	15.490
UNA	8	96	70.395	21.737
NEW_ASIAN_VALUES	0	1	0.209	0.412
POPUL	0.400	299.400	35.198	60.858
AREA	300	16,381,400	1,102,563	2,990,866
DENSITY	2	6,581	491.093	1,379.218
URBAN	33	100	70.837	17.152
LIFE	50	82	73.651	5.936
MORTALITY	2	90	16.674	17.590
GDP_PPP	2,501	122,884	25,359	24,152
GOV_EXP_ON_EDU	0	11	4.343	1.859
EMPHASIS_M_S	0	1	0.535	0.505
NURSERY	0	1	0.233	0.427
6LESS	0	1	0.860	0.351
POST_COMMUNIST	0	1	0.256	0.441
POL_STABILITY_I	-2.120	1.270	0.044	0.914
CORRUP	-1.010	2.240	0.304	0.854
ARGI	0.180	53.970	16.099	15.246
FEMALE_EMPLOYMENT	13.805	49.360	37.233	11.294
MIGRANT_STOCK	0.052	77.191	13.624	17.651
OIL	0	50.909	5.792	12.840
PROP_PRIM_PRIVATE	0.031	69.171	12.227	14.845
MEAN_Y_SCHOOL	5.648	13.256	9.757	2.331
ADVANCED	0	1	0.279	0.454

Table C.2: Correlation table of all 28 potential explanatory variables from our second-stage international analysis investigating international differences among countries

Independent variable	index_mod	EXIT_EXAM	PD	INDIV	MAS	UNA	NEW_ASIAN_VALUES	POPUL	AREA	DENSITY	URBAN	LIFE	MORTALITY	GDP_PPP	GOV_EXP_ON_EDU	EMPHASIS_M_S	NURSERY	6LESS	POST_COMMUNIST	POL_STABILITY_I	CORRUP	ARGI	FEMALE_EMPLOYMENT	MIGRANT_STOCK	OIL	PROP_PRIM_PRIVATE	MEAN_Y_SCHOOL	ADVANCED
index_mod	1																											
EXIT_EXAM	0.029	1																										
PD	0.26	-0.311	1																									
INDIV	-0.218	0.168	-0.743	1																								
MAS	-0.077	-0.071	-0.269	0.401	1																							
UNA	0.123	-0.273	0.226	-0.248	-0.192	1																						
NEW_ASIAN_VALUES	-0.451	0.124	0.107	-0.331	0.146	-0.398	1																					
POPUL	-0.008	0.175	-0.111	0.166	0.126	-0.113	0.199	1																				
AREA	0.02	0.135	-0.005	0.284	-0.02	0.013	-0.101	0.584	1																			
DENSITY	-0.221	0.042	0.011	-0.157	0.062	-0.545	0.439	-0.135	-0.126	1																		
URBAN	-0.173	-0.224	-0.217	0.308	0.273	-0.361	-0.046	-0.134	0.068	0.44	1																	
LIFE	-0.387	-0.148	-0.411	0.356	0.394	-0.244	0.167	-0.032	-0.102	0.331	0.535	1																
MORTALITY	0.483	0.056	0.316	-0.356	-0.192	0.122	-0.184	-0.039	-0.042	-0.223	-0.381	-0.882	1															
GDP_PPP	0.02	-0.346	0.042	0.111	0.155	-0.236	0.001	-0.041	0.016	0.289	0.668	0.455	-0.321	1														
GOV_EXP_ON_EDU	0.081	0.21	-0.03	0.03	-0.181	0.189	-0.168	-0.423	-0.183	-0.111	-0.148	-0.288	0.322	-0.177	1													
EMPHASIS_M_S	-0.049	-0.021	-0.177	0.063	-0.073	-0.151	0.021	-0.065	-0.037	-0.024	0.142	0.08	0.042	0.142	-0.034	1												
NURSERY	0.066	0.283	0.012	-0.068	0.066	0.216	-0.148	-0.167	-0.132	-0.145	-0.128	-0.183	0.134	-0.254	0.144	-0.259	1											
6LESS	0.14	-0.207	-0.091	0.164	0.383	-0.153	-0.123	-0.088	0.056	0.112	0.285	0.262	-0.062	0.263	-0.009	0.028	-0.255	1										
POST_COMMUNIST	-0.206	0.04	0.149	0.057	-0.233	0.404	-0.171	-0.121	0.127	-0.179	-0.334	-0.183	-0.056	-0.266	-0.001	-0.308	0.182	-0.225	1									
POL_STABILITY_I	-0.311	-0.012	-0.139	0.316	0.088	-0.132	0.161	-0.182	-0.097	0.264	0.303	0.266	-0.196	0.435	0.021	0.235	-0.24	0.03	0.134	1								
CORRUP	-0.293	0.036	-0.532	0.518	0.365	-0.53	0.161	-0.054	-0.056	0.483	0.553	0.525	-0.322	0.506	0.017	0.293	-0.229	0.354	-0.325	0.631	1							
ARGI	0.107	-0.019	0.321	-0.459	-0.227	0.199	0.174	0.058	-0.122	-0.261	-0.743	-0.59	0.553	-0.599	-0.028	-0.128	0.191	-0.256	0.223	-0.349	-0.58	1						
FEMALE_EMPLOYMENT	-0.47	0.385	-0.323	0.247	-0.044	-0.021	0.212	0.132	0.192	0.077	-0.236	-0.084	-0.064	-0.316	-0.005	-0.049	0.276	-0.252	0.437	0.272	0.218	0.209	1					
MIGRANT_STOCK	0.218	-0.328	0.16	-0.072	0.022	-0.221	-0.104	-0.235	-0.05	0.325	0.687	0.312	-0.189	0.82	-0.139	0.139	-0.161	0.21	-0.262	0.194	0.307	-0.502	-0.459	1				
OIL	0.33	-0.376	0.37	-0.191	-0.036	0.101	-0.17	-0.048	0.056	-0.138	0.24	0.05	0.007	0.566	-0.018	0.186	-0.219	0.152	-0.211	0.029	-0.138	-0.187	-0.612	0.496	1			
PROP_PRIM_PRIVATE	0.292	-0.24	0.009	0.003	0.142	-0.126	-0.158	-0.139	-0.068	0.01	0.467	0.11	0.033	0.372	-0.106	-0.098	0.09	0.177	-0.398	-0.201	0.015	-0.2	-0.487	0.577	0.264	1		
MEAN_Y_SCHOOL	-0.431	0.3	-0.498	0.656	0.243	-0.161	-0.065	0.097	0.276	0.061	0.268	0.336	-0.404	0.161	0.022	0.092	0.048	0.116	0.361	0.385	0.494	-0.388	0.491	0.022	-0.246	-0.244	1	
ADVANCED	-0.453	0.193	-0.532	0.377	0.286	-0.42	0.317	0.15	0.077	0.368	0.37	0.611	-0.451	0.268	-0.076	0.268	-0.097	0.251	-0.246	0.386	0.78	-0.4	0.384	0.071	-0.266	-0.158	0.501	1

Appendix D

In this part, we report descriptive statistics of modern and traditional teaching practices currently used in a classroom on a country level. Table D.1 shows an average frequency of the use of modern and traditional teaching methods per subject in a given country.

Table D.1: Descriptive statistics of teaching practices in a classroom

Country	Teaching practices		Mean	SD	Correlation (modern, traditional)
Armenia	Modern teaching practices	Total average	0.51	*0.09*	0.571
		Math	0.50	*0.08*	
		Biology	0.53	*0.09*	
		Chemistry	0.50	*0.10*	
		Earth science	0.48	*0.09*	
		Physics	0.54	*0.10*	
	Traditional teaching practices	Total average	0.66	*0.10*	
		Math	0.75	*0.09*	
		Biology	0.62	*0.10*	
		Chemistry	0.64	*0.11*	
		Earth science	0.59	*0.10*	
		Physics	0.68	*0.10*	
Australia	Modern teaching practices	Total average	0.47	*0.09*	0.358
		Math	0.45	*0.09*	
		Science	0.48	*0.09*	
	Traditional teaching practices	Total average	0.50	*0.09*	
		Math	0.52	*0.09*	
		Science	0.49	*0.08*	
Bahrain	Modern teaching practices	Total average	0.54	*0.09*	0.502
		Math	0.53	*0.08*	
		Science	0.56	*0.09*	
	Traditional teaching practices	Total average	0.65	*0.08*	
		Math	0.65	*0.07*	
		Science	0.65	*0.08*	
Bosnia and Herzegovina	Modern teaching practices	Total average	0.55	*0.10*	0.603
		Math	0.48	*0.11*	
		Biology	0.57	*0.09*	
		Chemistry	0.56	*0.11*	
		Earth science	0.55	*0.10*	
		Physics	0.56	*0.10*	
	Traditional teaching practices	Total average	0.62	*0.08*	
		Math	0.63	*0.07*	
		Biology	0.62	*0.08*	
		Chemistry	0.61	*0.08*	
		Earth science	0.61	*0.08*	
		Physics	0.62	*0.08*	
Botswana	Modern teaching practices	Total average	0.59	*0.06*	0.136

		Math	0.57	*0.06*	
		Science	0.61	*0.06*	
	Traditional teaching practices	Total average	0.49	*0.06*	
		Math	0.50	*0.06*	
		Science	0.48	*0.06*	
Chinese Taipei	Modern teaching practices	Total average	0.33	*0.07*	0.485
		Math	0.29	*0.06*	
		Science	0.37	*0.08*	
	Traditional teaching practices	Total average	0.55	*0.09*	
		Math	0.56	*0.08*	
		Science	0.54	*0.09*	
Colombia	Modern teaching practices	Total average	0.61	*0.08*	0.554
		Math	0.60	*0.08*	
		Science	0.61	*0.09*	
	Traditional teaching practices	Total average	0.66	*0.07*	
		Math	0.68	*0.07*	
		Science	0.64	*0.07*	
Cyprus	Modern teaching practices	Total average	0.53	*0.08*	0.491
		Math	0.51	*0.07*	
		Chemistry	0.56	*0.10*	
		Earth science	0.48	*0.10*	
		Physics	0.59	*0.11*	
	Traditional teaching practices	Total average	0.50	*0.07*	
		Math	0.54	*0.07*	
		Chemistry	0.50	*0.09*	
		Earth science	0.47	*0.09*	
		Physics	0.50	*0.09*	
Czech Republic	Modern teaching practices	Total average	0.44	*0.08*	0.456
		Math	0.40	*0.08*	
		Biology	0.45	*0.07*	
		Chemistry	0.49	*0.08*	
		Earth science	0.39	*0.08*	
		Physics	0.48	*0.09*	
	Traditional teaching practices	Total average	0.64	*0.08*	
		Math	0.67	*0.08*	
		Biology	0.61	*0.09*	
		Chemistry	0.68	*0.08*	
		Earth science	0.60	*0.10*	
		Physics	0.65	*0.09*	
Egypt	Modern teaching practices	Total average	0.62	*0.07*	0.550
		Math	0.62	*0.06*	
		Science	0.63	*0.08*	
	Traditional teaching practices	Total average	0.67	*0.07*	
		Math	0.66	*0.07*	
		Science	0.69	*0.06*	
El Salvador	Modern teaching practices	Total average	0.58	*0.08*	0.519
		Math	0.59	*0.08*	
		Science	0.58	*0.08*	
	Traditional teaching practices	Total average	0.67	*0.07*	
		Math	0.71	*0.07*	
		Science	0.63	*0.07*	
England	Modern teaching practices	Total average	0.46	*0.08*	0.410
		Math	0.44	*0.08*	
		Science	0.49	*0.07*	
	Traditional teaching practices	Total average	0.55	*0.09*	
		Math	0.54	*0.10*	

		Science	0.55	*0.08*	
Georgia	Modern teaching practices	Total average	0.50	*0.11*	0.775
		Math	0.55	*0.09*	
		Biology	0.50	*0.10*	
		Chemistry	0.47	*0.10*	
		Earth science	0.47	*0.14*	
		Physics	0.50	*0.10*	
	Traditional teaching practices	Total average	0.60	*0.08*	
		Math	0.65	*0.06*	
		Biology	0.60	*0.08*	
		Chemistry	0.57	*0.08*	
		Earth science	0.57	*0.12*	
		Physics	0.59	*0.08*	
Ghana	Modern teaching practices	Total average	0.65	*0.08*	0.563
		Math	0.63	*0.08*	
		Science	0.66	*0.08*	
	Traditional teaching practices	Total average	0.62	*0.09*	
		Math	0.58	*0.10*	
		Science	0.65	*0.07*	
Hong Kong SAR	Modern teaching practices	Total average	0.45	*0.07*	0.293
		Math	0.38	*0.06*	
		Science	0.52	*0.09*	
	Traditional teaching practices	Total average	0.49	*0.07*	
		Math	0.50	*0.07*	
		Science	0.47	*0.07*	
Hungary	Modern teaching practices	Total average	0.45	*0.09*	0.514
		Math	0.40	*0.08*	
		Biology	0.49	*0.08*	
		Chemistry	0.46	*0.10*	
		Earth science	0.44	*0.09*	
		Physics	0.48	*0.10*	
	Traditional teaching practices	Total average	0.55	*0.09*	
		Math	0.56	*0.08*	
		Biology	0.54	*0.09*	
		Chemistry	0.56	*0.09*	
		Earth science	0.53	*0.09*	
		Physics	0.57	*0.10*	
Indonesia	Modern teaching practices	Total average	0.47	*0.05*	0.587
		Math	0.47	*0.07*	
		Biology	0.49	*0.08*	
		Physics	0.46	*0.09*	
	Traditional teaching practices	Total average	0.64	*0.04*	
		Math	0.68	*0.07*	
		Biology	0.64	*0.07*	
		Physics	0.59	*0.07*	
Islamic Rep. of Iran	Modern teaching practices	Total average	0.56	*0.10*	0.142
		Math	0.55	*0.09*	
		Science	0.57	*0.11*	
	Traditional teaching practices	Total average	0.61	*0.08*	
		Math	0.61	*0.08*	
		Science	0.61	*0.08*	
Israel	Modern teaching practices	Total average	0.48	*0.09*	0.378
		Math	0.48	*0.08*	
		Science	0.47	*0.10*	
	Traditional teaching practices	Total average	0.65	*0.08*	
		Math	0.72	*0.07*	

		Science	0.59	*0.10*	
Japan	Modern teaching practices	Total average	0.52	*0.08*	0.074
		Math	0.61	*0.07*	
		Science	0.44	*0.08*	
	Traditional teaching practices	Total average	0.66	*0.07*	
		Math	0.63	*0.07*	
		Science	0.69	*0.07*	
Jordan	Modern teaching practices	Total average	0.65	*0.10*	0.640
		Math	0.63	*0.08*	
		Science	0.66	*0.11*	
	Traditional teaching practices	Total average	0.76	*0.08*	
		Math	0.77	*0.07*	
		Science	0.74	*0.10*	
Korea	Modern teaching practices	Total average	0.29	*0.05*	0.396
		Math	0.28	*0.05*	
		Science	0.30	*0.05*	
	Traditional teaching practices	Total average	0.52	*0.06*	
		Math	0.53	*0.06*	
		Science	0.51	*0.06*	
Rep. of Kuwait	Modern teaching practices	Total average	0.59	*0.08*	0.499
		Math	0.55	*0.08*	
		Science	0.63	*0.08*	
	Traditional teaching practices	Total average	0.69	*0.07*	
		Math	0.69	*0.07*	
		Science	0.70	*0.07*	
Lebanon	Modern teaching practices	Total average	0.55	*0.08*	0.395
		Math	0.52	*0.09*	
		Biology	0.56	*0.10*	
		Chemistry	0.56	*0.09*	
		Physics	0.56	*0.10*	
	Traditional teaching practices	Total average	0.62	*0.09*	
		Math	0.59	*0.10*	
		Biology	0.62	*0.11*	
		Chemistry	0.62	*0.11*	
		Physics	0.63	*0.11*	
Lithuania	Modern teaching practices	Total average	0.38	*0.09*	0.477
		Math	0.37	*0.09*	
		Biology	0.37	*0.09*	
		Chemistry	0.38	*0.08*	
		Earth science	0.38	*0.09*	
		Physics	0.41	*0.09*	
	Traditional teaching practices	Total average	0.63	*0.09*	
		Math	0.65	*0.09*	
		Biology	0.56	*0.09*	
		Chemistry	0.66	*0.09*	
		Earth science	0.58	*0.09*	
		Physics	0.68	*0.10*	
Malaysia	Modern teaching practices	Total average	0.47	*0.08*	0.629
		Math	0.46	*0.07*	
		Science	0.49	*0.09*	
	Traditional teaching practices	Total average	0.56	*0.09*	
		Math	0.58	*0.09*	
		Science	0.54	*0.09*	
Malta	Modern teaching practices	Total average	0.45	*0.13*	0.518
		Math	0.38	*0.09*	
		Biology	0.53	*0.16*	

		Chemistry	0.51	*0.19*	
		Earth science	0.35	*0.11*	
		Physics	0.48	*0.10*	
	Traditional teaching practices	Total average	0.46	*0.12*	
		Math	0.45	*0.08*	
		Biology	0.46	*0.15*	
		Chemistry	0.52	*0.18*	
		Earth science	0.37	*0.10*	
		Physics	0.50	*0.10*	
Mongolia	Modern teaching practices	Total average	0.50	*0.09*	0.771
		Math	0.54	*0.08*	
		Biology	0.48	*0.10*	
		Chemistry	0.49	*0.09*	
		Earth science	0.48	*0.09*	
		Physics	0.50	*0.09*	
	Traditional teaching practices	Total average	0.56	*0.08*	
		Math	0.59	*0.07*	
		Biology	0.55	*0.09*	
		Chemistry	0.56	*0.08*	
		Earth science	0.54	*0.08*	
		Physics	0.56	*0.08*	
Oman	Modern teaching practices	Total average	0.64	*0.09*	0.385
		Math	0.64	*0.08*	
		Science	0.65	*0.09*	
	Traditional teaching practices	Total average	0.64	*0.09*	
		Math	0.66	*0.09*	
		Science	0.63	*0.10*	
Palestinian Nat'l Auth.	Modern teaching practices	Total average	0.57	*0.09*	0.243
		Math	0.55	*0.09*	
		Science	0.60	*0.10*	
	Traditional teaching practices	Total average	0.60	*0.11*	
		Math	0.59	*0.11*	
		Science	0.60	*0.11*	
Qatar	Modern teaching practices	Total average	0.59	*0.09*	0.453
		Math	0.58	*0.08*	
		Science	0.60	*0.09*	
	Traditional teaching practices	Total average	0.62	*0.07*	
		Math	0.61	*0.08*	
		Science	0.62	*0.07*	
Romania	Modern teaching practices	Total average	0.49	*0.12*	0.555
		Math	0.46	*0.11*	
		Biology	0.49	*0.11*	
		Chemistry	0.50	*0.13*	
		Earth science	0.49	*0.12*	
		Physics	0.49	*0.12*	
	Traditional teaching practices	Total average	0.59	*0.12*	
		Math	0.68	*0.10*	
		Biology	0.55	*0.13*	
		Chemistry	0.58	*0.13*	
		Earth science	0.57	*0.13*	
		Physics	0.56	*0.13*	
Russian Federation	Modern teaching practices	Total average	0.53	*0.08*	0.523
		Math	0.49	*0.08*	
		Biology	0.52	*0.08*	
		Chemistry	0.55	*0.09*	

		Earth science	0.51	*0.09*	
		Physics	0.57	*0.09*	
	Traditional teaching practices	Total average	0.70	*0.08*	
		Math	0.73	*0.08*	
		Biology	0.60	*0.08*	
		Chemistry	0.79	*0.09*	
		Earth science	0.60	*0.09*	
		Physics	0.80	*0.09*	
Saudi Arabia	Modern teaching practices	Total average	0.53	*0.10*	0.583
		Math	0.52	*0.10*	
		Science	0.54	*0.10*	
	Traditional teaching practices	Total average	0.60	*0.10*	
		Math	0.59	*0.09*	
		Science	0.61	*0.10*	
Scotland	Modern teaching practices	Total average	0.51	*0.08*	0.320
		Math	0.46	*0.09*	
		Science	0.56	*0.08*	
	Traditional teaching practices	Total average	0.58	*0.08*	
		Math	0.59	*0.08*	
		Science	0.58	*0.08*	
Serbia	Modern teaching practices	Total average	0.46	*0.10*	0.413
		Math	0.40	*0.10*	
		Biology	0.52	*0.10*	
		Chemistry	0.46	*0.12*	
		Earth science	0.47	*0.10*	
		Physics	0.45	*0.11*	
	Traditional teaching practices	Total average	0.52	*0.09*	
		Math	0.55	*0.09*	
		Biology	0.52	*0.09*	
		Chemistry	0.53	*0.09*	
		Earth science	0.49	*0.09*	
		Physics	0.54	*0.08*	
Singapore	Modern teaching practices	Total average	0.47	*0.08*	0.523
		Math	0.44	*0.08*	
		Science	0.50	*0.09*	
	Traditional teaching practices	Total average	0.53	*0.08*	
		Math	0.53	*0.07*	
		Science	0.53	*0.09*	
Slovenia	Modern teaching practices	Total average	0.50	*0.07*	0.452
		Math	0.46	*0.08*	
		Biology	0.51	*0.09*	
		Chemistry	0.53	*0.09*	
		Physics	0.51	*0.09*	
	Traditional teaching practices	Total average	0.66	*0.06*	
		Math	0.69	*0.07*	
		Biology	0.66	*0.07*	
		Chemistry	0.66	*0.07*	
		Physics	0.65	*0.09*	
Thailand	Modern teaching practices	Total average	0.55	*0.07*	0.707
		Math	0.53	*0.06*	
		Science	0.57	*0.09*	
	Traditional teaching practices	Total average	0.59	*0.07*	
		Math	0.60	*0.07*	
		Science	0.58	*0.07*	
Tunisia	Modern teaching practices	Total average	0.58	*0.08*	0.234

		Math	0.51	*0.08*	
		Science	0.66	*0.09*	
	Traditional teaching practices	Total average	0.63	*0.09*	
		Math	0.63	*0.09*	
		Science	0.63	*0.09*	
Turkey	Modern teaching practices	Total average	0.55	*0.08*	0.536
		Math	0.55	*0.07*	
		Science	0.55	*0.10*	
	Traditional teaching practices	Total average	0.55	*0.08*	
		Math	0.58	*0.08*	
		Science	0.52	*0.08*	
Ukraine	Modern teaching practices	Total average	0.55	*0.08*	0.505
		Math	0.52	*0.08*	
		Biology	0.56	*0.08*	
		Chemistry	0.57	*0.08*	
		Earth science	0.54	*0.08*	
		Physics	0.58	*0.08*	
	Traditional teaching practices	Total average	0.73	*0.08*	
		Math	0.73	*0.08*	
		Biology	0.71	*0.08*	
		Chemistry	0.76	*0.09*	
		Earth science	0.70	*0.08*	
		Physics	0.74	*0.09*	
United States	Modern teaching practices	Total average	0.53	*0.11*	0.208
		Math	0.52	*0.11*	
		Science	0.54	*0.11*	
	Traditional teaching practices	Total average	0.63	*0.09*	
		Math	0.66	*0.10*	
		Science	0.60	*0.09*	

Note: Students weights are used. Standard deviation (SD) in italics. The class-aggregated modern and traditional teaching practices index expresses a proportion of a particular class in which students experience modern or traditional teaching practices.

Appendix E

Tables E.1 to E.43 below report teacher and class summary statistics on country-level. There are 43 separate tables with teacher and class descriptive statistics, one for each country.

Table E.1: Descriptive statistics: Teacher and class characteristics in Armenia

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		240 Teachers		189 Teachers		161 Teachers		178 Teachers		209 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.03	*0.16*	0.04	*0.19*	0.06	*0.24*	0.03	*0.16*	0.01	*0.08*
	25-29	0.09	*0.29*	0.06	*0.23*	0.12	*0.32*	0.07	*0.26*	0.01	*0.11*
	30-39	0.24	*0.43*	0.27	*0.44*	0.29	*0.45*	0.31	*0.46*	0.25	*0.43*
	40-49	0.28	*0.45*	0.34	*0.47*	0.37	*0.48*	0.24	*0.43*	0.38	*0.48*
	50-59	0.32	*0.47*	0.27	*0.45*	0.13	*0.34*	0.31	*0.46*	0.32	*0.47*
	Over 60	0.04	*0.19*	0.02	*0.13*	0.03	*0.18*	0.04	*0.19*	0.03	*0.16*
	Female	0.82	*0.39*	0.94	*0.23*	0.82	*0.38*	0.87	*0.33*	0.80	*0.40*
Experience	1-2 years	0.02	*0.15*	0.02	*0.14*	0.05	*0.22*	0.03	*0.16*	0.01	*0.08*
	3-5 years	0.06	*0.24*	0.05	*0.22*	0.06	*0.23*	0.03	*0.16*	0.03	*0.18*
	Over 6 years	0.92	*0.27*	0.93	*0.25*	0.89	*0.31*	0.95	*0.22*	0.96	*0.19*
	University diploma	0.98	*0.14*	0.98	*0.13*	0.94	*0.23*	0.99	*0.10*	0.98	*0.14*
Motivation on 1-4 scale	Discusses concepts with others	2.44	*1.01*	2.40	*0.81*	2.42	*0.83*	2.51	*0.81*	2.38	*0.86*
	Prepares material	2.64	*0.87*	2.45	*0.77*	2.55	*0.74*	2.60	*0.77*	2.65	*0.91*
	Visits other classes	2.61	*0.80*	2.51	*0.70*	2.52	*0.73*	2.55	*0.75*	2.55	*0.72*
	Informal visits	2.64	*0.90*	2.72	*0.81*	2.62	*0.92*	2.68	*0.80*	2.69	*0.84*
Further development	Subject content course	0.68	*0.47*	0.60	*0.49*	0.36	*0.48*	0.52	*0.50*	0.00	*0.00*
	Holds a certificate	0.96	*0.19*	0.98	*0.14*	0.93	*0.26*	0.97	*0.18*	0.91	*0.28*
	Curriculum improvement course	0.68	*0.47*	0.54	*0.50*	0.35	*0.48*	0.61	*0.49*	0.47	*0.50*
	Subject related to IT	0.35	*0.48*	0.29	*0.45*	0.29	*0.45*	0.37	*0.48*	0.23	*0.42*
	Critical thinking course	0.36	*0.48*	0.47	*0.50*	0.44	*0.50*	0.43	*0.49*	0.35	*0.48*
Class characteristics	Class size	22.57	*7.25*	22.74	*7.49*	22.60	*7.45*	22.81	*7.60*	22.40	*7.05*
	Minutes in class	186.82	*82.33*	127.78	*66.18*	107.77	*57.44*	137.92	*74.13*	110.90	*45.83*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.2: Descriptive statistics: Teacher and class characteristics in Australia

		Mathematics		Science	
		221 Teachers		466 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.08	*0.27*	0.14	*0.35*
	25-29	0.14	*0.35*	0.14	*0.35*
	30-39	0.28	*0.45*	0.21	*0.41*
	40-49	0.22	*0.42*	0.25	*0.44*
	50-59	0.26	*0.44*	0.21	*0.41*
	Over 60	0.02	*0.12*	0.04	*0.19*
	Female	0.49	*0.50*	0.44	*0.50*
Experience	1-2 years	0.15	*0.35*	0.28	*0.45*
	3-5 years	0.13	*0.34*	0.11	*0.32*
	Over 6 years	0.72	*0.45*	0.61	*0.49*
	University diploma	1.00	*0.00*	0.93	*0.26*
Motivation on 1-4 scale	Discusses concepts with others	2.38	*0.77*	2.48	*0.95*
	Prepares material	2.29	*0.85*	2.37	*1.00*
	Visits other classes	1.21	*0.53*	1.27	*0.49*
	Informal visits	1.34	*0.67*	1.35	*0.58*
Further development	Subject content course	0.64	*0.48*	0.50	*0.50*
	Holds a certificate	0.98	*0.15*	0.88	*0.32*
	Curriculum improvement course	0.68	*0.47*	0.52	*0.50*
	Subject related to IT	0.55	*0.50*	0.52	*0.50*
	Critical thinking course	0.46	*0.50*	0.47	*0.50*
Class characteristics	Class size	24.64	*5.05*	20.70	*10.16*
	Minutes in class	202.16	*38.76*	164.04	*73.02*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.3: Descriptive statistics: Teacher and class characteristics in Bahrain

		Mathematics		Science	
		115 Teachers		133 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.03	*0.18*	0.06	*0.23*
	25-29	0.13	*0.34*	0.15	*0.36*
	30-39	0.53	*0.50*	0.53	*0.50*
	40-49	0.27	*0.44*	0.25	*0.43*
	50-59	0.04	*0.19*	0.01	*0.11*
	Over 60	0.00	*0.06*	0.00	*0.07*
	Female	0.48	*0.50*	0.49	*0.50*
Experience	1-2 years	0.12	*0.33*	0.18	*0.39*
	3-5 years	0.13	*0.34*	0.14	*0.35*
	Over 6 years	0.74	*0.44*	0.67	*0.47*
	University diploma	0.96	*0.19*	0.95	*0.21*
Motivation on 1-4 scale	Discusses concepts with others	2.58	*0.90*	2.45	*0.86*
	Prepares material	2.66	*0.93*	2.61	*0.92*
	Visits other classes	1.58	*0.62*	1.57	*0.54*
	Informal visits	1.53	*0.74*	1.52	*0.66*
Further development	Subject content course	0.45	*0.50*	0.49	*0.50*
	Holds a certificate	0.96	*0.19*	0.89	*0.31*
	Curriculum improvement course	0.26	*0.44*	0.33	*0.47*
	Subject related to IT	0.66	*0.48*	0.63	*0.48*
	Critical thinking course	0.52	*0.50*	0.57	*0.50*
Class characteristics	Class size	30.04	*4.28*	25.58	*11.67*
	Minutes in class	155.72	*94.13*	111.81	*86.46*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.4: Descriptive statistics: Teacher and class characteristics in Bosnia and Herzegovina

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		156 Teachers		167 Teachers		169 Teachers		161 Teachers		161 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.01	*0.08*	0.01	*0.11*	0.05	*0.22*	0.03	*0.18*	0.02	*0.15*
	25-29	0.06	*0.24*	0.11	*0.32*	0.10	*0.30*	0.06	*0.23*	0.07	*0.25*
	30-39	0.20	*0.40*	0.17	*0.37*	0.23	*0.42*	0.14	*0.35*	0.26	*0.44*
	40-49	0.21	*0.41*	0.19	*0.39*	0.26	*0.44*	0.28	*0.45*	0.22	*0.41*
	50-59	0.39	*0.49*	0.36	*0.48*	0.24	*0.43*	0.36	*0.48*	0.30	*0.46*
	Over 60	0.13	*0.34*	0.16	*0.37*	0.11	*0.31*	0.13	*0.33*	0.13	*0.34*
	Female	0.53	*0.50*	0.64	*0.48*	0.54	*0.50*	0.66	*0.47*	0.56	*0.50*
Experience	1-2 years	0.09	*0.28*	0.09	*0.29*	0.17	*0.38*	0.09	*0.28*	0.09	*0.28*
	3-5 years	0.05	*0.21*	0.09	*0.29*	0.08	*0.27*	0.04	*0.20*	0.08	*0.27*
	Over 6 years	0.86	*0.34*	0.82	*0.39*	0.75	*0.43*	0.87	*0.33*	0.84	*0.37*
	University diploma	0.97	*0.17*	0.00	*0.00*	0.00	*0.00*	0.90	*0.30*	0.93	*0.25*
Motivation on 1-4 scale	Discusses concepts with others	2.61	*0.91*	2.62	*0.87*	2.70	*1.00*	2.60	*0.92*	2.58	*0.90*
	Prepares material	2.49	*0.95*	2.52	*0.95*	2.64	*0.97*	2.54	*0.95*	2.57	*0.92*
	Visits other classes	1.44	*0.57*	1.55	*0.63*	1.68	*0.74*	1.66	*0.72*	1.44	*0.66*
	Informal visits	1.40	*0.60*	1.52	*0.65*	1.64	*0.75*	1.63	*0.79*	1.46	*0.68*
Further development	Subject content course	0.59	*0.49*	0.37	*0.48*	0.32	*0.47*	0.39	*0.49*	0.00	*0.00*
	Holds a certificate	0.97	*0.17*	0.96	*0.19*	0.94	*0.23*	0.97	*0.17*	0.97	*0.18*
	Curriculum improvement course	0.56	*0.50*	0.46	*0.50*	0.36	*0.48*	0.46	*0.50*	0.45	*0.50*
	Subject related to IT	0.36	*0.48*	0.38	*0.49*	0.39	*0.49*	0.42	*0.49*	0.47	*0.50*
	Critical thinking course	0.45	*0.50*	0.42	*0.49*	0.47	*0.50*	0.49	*0.50*	0.42	*0.49*
Class characteristics	Class size	22.67	*6.11*	21.72	*7.78*	21.09	*8.20*	21.54	*8.77*	21.70	*7.58*
	Minutes in class	173.03	*22.24*	97.59	*39.25*	80.31	*26.58*	93.99	*39.52*	91.31	*31.46*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.5: Descriptive statistics: Teacher and class characteristics in Bostwana

		Mathematics		Science	
		128 Teachers		150 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.04	*0.21*	0.05	*0.21*
	25-29	0.27	*0.44*	0.38	*0.48*
	30-39	0.59	*0.49*	0.52	*0.50*
	40-49	0.08	*0.26*	0.04	*0.20*
	50-59	0.01	*0.11*	0.01	*0.08*
	Over 60	0.01	*0.07*	0.01	*0.08*
	Female	0.46	*0.50*	0.41	*0.49*
Experience	1-2 years	0.21	*0.41*	0.20	*0.40*
	3-5 years	0.20	*0.40*	0.27	*0.44*
	Over 6 years	0.60	*0.49*	0.54	*0.50*
	University diploma	0.97	*0.16*	0.97	*0.16*
Motivation on 1-4 scale	Discusses concepts with others	2.46	*0.90*	2.54	*0.90*
	Prepares material	2.47	*0.99*	2.78	*1.02*
	Visits other classes	1.61	*0.60*	1.69	*0.75*
	Informal visits	1.59	*0.63*	1.65	*0.66*
Further development	Subject content course	0.11	*0.31*	0.14	*0.35*
	Holds a certificate	0.99	*0.12*	0.93	*0.26*
	Curriculum improvement course	0.10	*0.30*	0.13	*0.33*
	Subject related to IT	0.11	*0.32*	0.17	*0.37*
	Critical thinking course	0.29	*0.45*	0.24	*0.43*
Class characteristics	Class size	36.95	*4.84*	34.22	*10.59*
	Minutes in class	226.27	*35.82*	193.46	*84.83*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.6: Descriptive statistics: Teacher and class characteristics in Chinese Tapei

		Mathematics		Science	
		145 Teachers		154 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.03	*0.18*	0.05	*0.21*
	25-29	0.14	*0.34*	0.14	*0.35*
	30-39	0.44	*0.50*	0.39	*0.49*
	40-49	0.29	*0.45*	0.34	*0.47*
	50-59	0.09	*0.29*	0.09	*0.29*
	Over 60	0.01	*0.11*	0.00	*0.00*
	Female	0.55	*0.50*	0.35	*0.48*
Experience	1-2 years	0.13	*0.33*	0.18	*0.38*
	3-5 years	0.18	*0.39*	0.16	*0.36*
	Over 6 years	0.69	*0.46*	0.66	*0.47*
	University diploma	0.95	*0.21*	0.86	*0.35*
Motivation on 1-4 scale	Discusses concepts with others	2.38	*0.74*	2.44	*0.80*
	Prepares material	1.73	*0.73*	1.70	*0.84*
	Visits other classes	1.36	*0.61*	1.34	*0.57*
	Informal visits	1.14	*0.39*	1.15	*0.41*
Further development	Subject content course	0.80	*0.40*	0.70	*0.46*
	Holds a certificate	0.95	*0.22*	0.94	*0.24*
	Curriculum improvement course	0.86	*0.34*	0.80	*0.40*
	Subject related to IT	0.74	*0.44*	0.73	*0.44*
	Critical thinking course	0.43	*0.49*	0.39	*0.49*
Class characteristics	Class size	34.30	*6.61*	32.69	*9.57*
	Minutes in class	236.49	*52.53*	210.04	*49.67*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.7: Descriptive statistics: Teacher and class characteristics in Colombia

		Mathematics		Science	
		128 Teachers		149 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.04	*0.20*	0.08	*0.27*
	25-29	0.23	*0.42*	0.11	*0.31*
	30-39	0.27	*0.45*	0.28	*0.45*
	40-49	0.22	*0.42*	0.29	*0.45*
	50-59	0.19	*0.39*	0.23	*0.42*
	Over 60	0.05	*0.21*	0.01	*0.09*
	Female	0.41	*0.49*	0.62	*0.49*
Experience	1-2 years	0.12	*0.32*	0.19	*0.40*
	3-5 years	0.13	*0.33*	0.06	*0.25*
	Over 6 years	0.76	*0.43*	0.74	*0.44*
	University diploma	0.89	*0.31*	0.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.58	*0.98*	2.26	*1.01*
	Prepares material	2.54	*0.99*	2.50	*1.09*
	Visits other classes	1.38	*0.78*	1.35	*0.70*
	Informal visits	1.49	*0.84*	1.69	*0.98*
Further development	Subject content course	0.63	*0.48*	0.64	*0.48*
	Holds a certificate	0.00	*0.00*	0.00	*0.00*
	Curriculum improvement course	0.68	*0.47*	0.67	*0.47*
	Subject related to IT	0.48	*0.50*	0.37	*0.48*
	Critical thinking course	0.57	*0.49*	0.44	*0.50*
Class characteristics	Class size	33.78	*8.97*	32.07	*14.81*
	Minutes in class	226.84	*93.52*	160.80	*101.07*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.8: Descriptive statistics: Teacher and class characteristics in Cyprus

		Mathematics		Earth Science		Chemistry		Physics	
		171 Teachers		114 Teachers		101 Teachers		117 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.15*	0.17	*0.37*	0.13	*0.34*	0.16	*0.37*
	25-29	0.06	*0.23*	0.04	*0.19*	0.00	*0.00*	0.29	*0.45*
	30-39	0.32	*0.47*	0.16	*0.37*	0.02	*0.13*	0.21	*0.41*
	40-49	0.35	*0.48*	0.34	*0.47*	0.60	*0.49*	0.25	*0.43*
	50-59	0.23	*0.42*	0.27	*0.44*	0.25	*0.43*	0.09	*0.28*
	Over 60	0.02	*0.13*	0.03	*0.16*	0.00	*0.00*	0.01	*0.08*
	Female	0.69	*0.46*	0.52	*0.50*	0.60	*0.49*	0.46	*0.50*
Experience	1-2 years	0.16	*0.36*	0.28	*0.45*	0.25	*0.43*	0.37	*0.48*
	3-5 years	0.15	*0.36*	0.15	*0.36*	0.15	*0.36*	0.19	*0.39*
	Over 6 years	0.69	*0.46*	0.57	*0.50*	0.59	*0.49*	0.44	*0.50*
	University diploma	0.99	*0.08*	0.84	*0.37*	0.87	*0.33*	0.83	*0.38*
Motivation on 1-4 scale	Discusses concepts with others	2.90	*0.96*	2.23	*0.95*	2.49	*0.92*	2.37	*0.95*
	Prepares material	2.83	*0.94*	2.17	*1.05*	2.53	*1.00*	2.33	*1.02*
	Visits other classes	1.26	*0.59*	1.15	*0.50*	1.13	*0.40*	1.15	*0.48*
	Informal visits	1.79	*0.96*	1.70	*0.92*	1.47	*0.80*	1.62	*0.88*
Further development	Subject content course	0.67	*0.47*	0.47	*0.50*	0.57	*0.50*	0.00	*0.00*
	Holds a certificate	0.96	*0.20*	0.75	*0.43*	0.80	*0.40*	0.75	*0.43*
	Curriculum improvement course	0.54	*0.50*	0.42	*0.49*	0.49	*0.50*	0.45	*0.50*
	Subject related to IT	0.57	*0.49*	0.42	*0.49*	0.59	*0.49*	0.69	*0.46*
	Critical thinking course	0.45	*0.50*	0.36	*0.48*	0.40	*0.49*	0.45	*0.50*
Class characteristics	Class size	23.93	*3.28*	17.85	*10.52*	19.77	*9.48*	19.37	*9.66*
	Minutes in class	129.71	*7.79*	62.97	*41.24*	43.08	*34.71*	72.26	*39.25*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.9: Descriptive statistics: Teacher and class characteristics in the Czech Republic

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		203	Teachers	202	Teachers	189	Teachers	194	Teachers	197	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.01	*0.12*	0.01	*0.08*	0.03	*0.16*	0.03	*0.16*	0.01	*0.10*
	25-29	0.12	*0.32*	0.14	*0.35*	0.22	*0.42*	0.22	*0.42*	0.09	*0.28*
	30-39	0.19	*0.40*	0.13	*0.33*	0.21	*0.41*	0.21	*0.41*	0.22	*0.41*
	40-49	0.33	*0.47*	0.22	*0.42*	0.22	*0.41*	0.22	*0.41*	0.34	*0.47*
	50-59	0.27	*0.44*	0.37	*0.48*	0.24	*0.43*	0.24	*0.43*	0.22	*0.41*
	Over 60	0.08	*0.27*	0.12	*0.33*	0.08	*0.27*	0.08	*0.27*	0.13	*0.34*
	Female	0.79	*0.40*	0.82	*0.38*	0.63	*0.48*	0.63	*0.48*	0.57	*0.50*
Experience	1-2 years	0.06	*0.23*	0.09	*0.29*	0.14	*0.35*	0.14	*0.35*	0.07	*0.26*
	3-5 years	0.10	*0.30*	0.07	*0.26*	0.11	*0.31*	0.11	*0.31*	0.09	*0.29*
	Over 6 years	0.84	*0.36*	0.83	*0.37*	0.76	*0.43*	0.76	*0.43*	0.83	*0.37*
	University diploma	0.97	*0.18*	0.97	*0.17*	0.95	*0.21*	0.95	*0.21*	0.96	*0.20*
Motivation on 1-4 scale	Discusses concepts with others	1.83	*0.71*	1.75	*0.69*	1.74	*0.74*	1.74	*0.74*	1.87	*0.76*
	Prepares material	1.97	*0.81*	1.90	*0.82*	1.93	*0.85*	1.93	*0.85*	2.06	*0.97*
	Visits other classes	1.15	*0.39*	1.11	*0.32*	1.15	*0.42*	1.15	*0.42*	1.15	*0.36*
	Informal visits	1.12	*0.35*	1.12	*0.35*	1.12	*0.38*	1.12	*0.38*	1.14	*0.35*
Further development	Subject content course	0.46	*0.50*	0.38	*0.48*	0.27	*0.45*	0.27	*0.45*	0.00	*0.00*
	Holds a certificate	0.96	*0.19*	0.94	*0.23*	0.94	*0.24*	0.94	*0.24*	0.95	*0.21*
	Curriculum improvement course	0.35	*0.48*	0.31	*0.46*	0.19	*0.40*	0.19	*0.40*	0.31	*0.46*
	Subject related to IT	0.47	*0.50*	0.57	*0.50*	0.52	*0.50*	0.52	*0.50*	0.56	*0.50*
	Critical thinking course	0.28	*0.45*	0.31	*0.46*	0.25	*0.43*	0.25	*0.43*	0.33	*0.47*
Class characteristics	Class size	23.68	*4.35*	23.75	*4.30*	23.84	*4.30*	23.84	*4.30*	23.61	*4.32*
	Minutes in class	198.66	*31.30*	88.54	*17.12*	82.68	*21.74*	82.68	*21.74*	88.67	*13.41*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.10: Descriptive statistics: Teacher and class characteristics in Egypt

		Mathematics		Science	
		220 Teachers		234 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.10	*0.30*	0.04	*0.20*
	25-29	0.07	*0.26*	0.18	*0.39*
	30-39	0.47	*0.50*	0.38	*0.48*
	40-49	0.31	*0.46*	0.37	*0.48*
	50-59	0.04	*0.21*	0.03	*0.16*
	Over 60	0.00	*0.02*	0.00	*0.00*
	Female	0.21	*0.41*	0.41	*0.49*
Experience	1-2 years	0.14	*0.34*	0.08	*0.27*
	3-5 years	0.07	*0.26*	0.14	*0.35*
	Over 6 years	0.79	*0.41*	0.78	*0.41*
	University diploma	0.89	*0.32*	0.95	*0.22*
Motivation on 1-4 scale	Discusses concepts with others	2.79	*0.78*	2.96	*0.85*
	Prepares material	2.63	*0.94*	3.07	*0.98*
	Visits other classes	2.01	*0.81*	2.10	*0.74*
	Informal visits	1.69	*0.73*	1.91	*0.80*
Further development	Subject content course	0.65	*0.48*	0.70	*0.46*
	Holds a certificate	0.67	*0.47*	0.61	*0.49*
	Curriculum improvement course	0.32	*0.47*	0.39	*0.49*
	Subject related to IT	0.53	*0.50*	0.48	*0.50*
	Critical thinking course	0.76	*0.43*	0.70	*0.46*
Class characteristics	Class size	37.15	*8.14*	35.42	*11.22*
	Minutes in class	154.03	*84.54*	138.96	*73.54*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.11: Descriptive statistics: Teacher and class characteristics in El Salvador

		Mathematics		Science	
		139 Teachers		147 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.05	*0.22*	0.05	*0.22*
	25-29	0.19	*0.39*	0.13	*0.34*
	30-39	0.48	*0.50*	0.52	*0.50*
	40-49	0.23	*0.42*	0.24	*0.43*
	50-59	0.05	*0.22*	0.05	*0.21*
	Over 60	0.00	*0.07*	0.01	*0.10*
	Female	0.59	*0.49*	0.49	*0.50*
Experience	1-2 years	0.09	*0.28*	0.09	*0.29*
	3-5 years	0.15	*0.36*	0.15	*0.36*
	Over 6 years	0.76	*0.43*	0.75	*0.43*
	University diploma	0.76	*0.43*	0.74	*0.44*
Motivation on 1-4 scale	Discusses concepts with others	1.71	*0.97*	1.93	*1.03*
	Prepares material	2.09	*1.11*	2.15	*1.13*
	Visits other classes	1.54	*0.86*	1.65	*1.00*
	Informal visits	1.71	*1.02*	1.67	*0.99*
Further development	Subject content course	0.40	*0.49*	0.37	*0.48*
	Holds a certificate	0.94	*0.23*	0.96	*0.20*
	Curriculum improvement course	0.22	*0.41*	0.23	*0.42*
	Subject related to IT	0.24	*0.43*	0.23	*0.42*
	Critical thinking course	0.43	*0.50*	0.44	*0.50*
Class characteristics	Class size	24.85	*10.53*	24.86	*11.01*
	Minutes in class	213.08	*42.51*	197.04	*61.25*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.12: Descriptive statistics: Teacher and class characteristics in England

		Mathematics		Science	
		211 Teachers		615 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.04	*0.19*	0.16	*0.37*
	25-29	0.15	*0.36*	0.18	*0.38*
	30-39	0.24	*0.43*	0.26	*0.44*
	40-49	0.26	*0.44*	0.20	*0.40*
	50-59	0.29	*0.46*	0.18	*0.39*
	Over 60	0.02	*0.14*	0.02	*0.15*
	Female	0.53	*0.50*	0.49	*0.50*
Experience	1-2 years	0.19	*0.39*	0.29	*0.46*
	3-5 years	0.20	*0.40*	0.17	*0.38*
	Over 6 years	0.61	*0.49*	0.53	*0.50*
	University diploma	0.98	*0.13*	0.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.40	*0.79*	2.43	*0.94*
	Prepares material	2.09	*0.85*	2.17	*0.92*
	Visits other classes	1.39	*0.56*	1.44	*0.66*
	Informal visits	1.52	*0.70*	1.49	*0.71*
Further development	Subject content course	0.79	*0.41*	0.70	*0.46*
	Holds a certificate	0.97	*0.18*	0.85	*0.36*
	Curriculum improvement course	0.61	*0.49*	0.65	*0.48*
	Subject related to IT	0.60	*0.49*	0.39	*0.49*
	Critical thinking course	0.39	*0.49*	0.43	*0.49*
Class characteristics	Class size	24.62	*7.52*	19.05	*12.83*
	Minutes in class	182.42	*32.59*	127.17	*73.40*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.13: Descriptive statistics: Teacher and class characteristics in Georgia

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		172 Teachers		158 Teachers		151 Teachers		144 Teachers		163 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.13*	0.06	*0.23*	0.06	*0.24*	0.12	*0.32*	0.07	*0.25*
	25-29	0.03	*0.16*	0.04	*0.19*	0.04	*0.19*	0.05	*0.21*	0.02	*0.14*
	30-39	0.23	*0.42*	0.21	*0.41*	0.28	*0.45*	0.17	*0.37*	0.24	*0.43*
	40-49	0.27	*0.44*	0.22	*0.41*	0.29	*0.45*	0.27	*0.45*	0.34	*0.47*
	50-59	0.25	*0.43*	0.27	*0.44*	0.25	*0.43*	0.28	*0.45*	0.17	*0.38*
	Over 60	0.21	*0.41*	0.21	*0.41*	0.09	*0.28*	0.11	*0.31*	0.15	*0.36*
	Female	0.87	*0.34*	0.89	*0.32*	0.83	*0.37*	0.90	*0.31*	0.79	*0.41*
Experience	1-2 years	0.04	*0.19*	0.07	*0.25*	0.09	*0.28*	0.14	*0.35*	0.13	*0.33*
	3-5 years	0.06	*0.24*	0.02	*0.15*	0.04	*0.18*	0.10	*0.30*	0.06	*0.23*
	Over 6 years	0.91	*0.29*	0.91	*0.29*	0.88	*0.33*	0.76	*0.43*	0.82	*0.39*
	University diploma	1.00	*0.00*	0.94	*0.24*	0.00	*0.00*	0.89	*0.32*	0.95	*0.22*
Motivation on 1-4 scale	Discusses concepts with others	2.61	*0.79*	2.76	*0.86*	2.76	*0.96*	2.45	*0.86*	2.54	*0.90*
	Prepares material	2.69	*1.02*	2.99	*0.96*	2.87	*1.03*	2.62	*1.03*	2.83	*1.01*
	Visits other classes	1.80	*0.49*	1.86	*0.50*	1.90	*0.60*	1.84	*0.68*	1.83	*0.56*
	Informal visits	1.74	*0.63*	1.72	*0.58*	1.75	*0.65*	1.74	*0.71*	1.70	*0.60*
Further development	Subject content course	0.48	*0.50*	0.52	*0.50*	0.39	*0.49*	0.42	*0.49*	0.00	*0.00*
	Holds a certificate	0.94	*0.24*	0.91	*0.28*	0.91	*0.29*	0.81	*0.39*	0.86	*0.34*
	Curriculum improvement course	0.51	*0.50*	0.62	*0.48*	0.44	*0.50*	0.54	*0.50*	0.55	*0.50*
	Subject related to IT	0.23	*0.42*	0.38	*0.48*	0.31	*0.46*	0.24	*0.43*	0.38	*0.48*
	Critical thinking course	0.58	*0.49*	0.59	*0.49*	0.62	*0.49*	0.66	*0.48*	0.61	*0.49*
Class characteristics	Class size	19.57	*8.73*	17.25	*9.78*	17.92	*10.73*	17.93	*11.38*	17.72	*9.82*
	Minutes in class	184.29	*15.44*	126.51	*77.28*	96.51	*66.45*	99.53	*70.22*	111.81	*73.71*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.14: Descriptive statistics: Teacher and class characteristics in Ghana

		Mathematics		Science	
		154 Teachers		171 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.20	*0.40*	0.18	*0.39*
	25-29	0.33	*0.47*	0.37	*0.48*
	30-39	0.25	*0.43*	0.31	*0.46*
	40-49	0.16	*0.37*	0.08	*0.27*
	50-59	0.04	*0.20*	0.06	*0.24*
	Over 60	0.01	*0.10*	0.00	*0.00*
	Female	0.09	*0.29*	0.07	*0.26*
Experience	1-2 years	0.30	*0.46*	0.29	*0.45*
	3-5 years	0.29	*0.46*	0.37	*0.48*
	Over 6 years	0.41	*0.49*	0.34	*0.47*
	University diploma	0.27	*0.44*	0.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.33	*0.91*	2.24	*0.93*
	Prepares material	2.53	*0.96*	2.59	*1.03*
	Visits other classes	2.41	*0.93*	2.25	*0.95*
	Informal visits	2.42	*0.97*	2.38	*0.91*
Further development	Subject content course	0.35	*0.48*	0.47	*0.50*
	Holds a certificate	0.71	*0.45*	0.75	*0.43*
	Curriculum improvement course	0.41	*0.49*	0.52	*0.50*
	Subject related to IT	0.13	*0.34*	0.22	*0.42*
	Critical thinking course	0.43	*0.50*	0.47	*0.50*
Class characteristics	Class size	37.75	*21.03*	34.68	*21.58*
	Minutes in class	224.39	*87.75*	171.44	*81.25*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.15: Descriptive statistics: Teacher and class characteristics in Hong Kong, SAR

		Mathematics		**Science**	
		138 Teachers		123 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.10	*0.30*	0.08	*0.28*
	25-29	0.18	*0.39*	0.16	*0.37*
	30-39	0.33	*0.47*	0.38	*0.49*
	40-49	0.27	*0.44*	0.25	*0.44*
	50-59	0.10	*0.29*	0.12	*0.32*
	Over 60	0.01	*0.11*	0.00	*0.04*
	Female	0.42	*0.49*	0.36	*0.48*
Experience	1-2 years	0.27	*0.45*	0.20	*0.40*
	3-5 years	0.14	*0.34*	0.09	*0.28*
	Over 6 years	0.59	*0.49*	0.72	*0.45*
	University diploma	0.94	*0.23*	0.96	*0.19*
Motivation on 1-4 scale	Discusses concepts with others	2.29	*0.67*	2.18	*0.69*
	Prepares material	1.93	*0.72*	1.91	*0.74*
	Visits other classes	1.46	*0.55*	1.47	*0.55*
	Informal visits	1.34	*0.53*	1.41	*0.56*
Further development	Subject content course	0.69	*0.46*	0.74	*0.44*
	Holds a certificate	0.94	*0.23*	0.89	*0.31*
	Curriculum improvement course	0.69	*0.46*	0.72	*0.45*
	Subject related to IT	0.61	*0.49*	0.54	*0.50*
	Critical thinking course	0.56	*0.50*	0.66	*0.47*
Class characteristics	Class size	34.33	*10.71*	36.92	*9.11*
	Minutes in class	243.93	*63.70*	148.90	*61.64*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.16: Descriptive statistics: Teacher and class characteristics in Hungary

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		227	Teachers	179	Teachers	183	Teachers	165	Teachers	172	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.03	*0.17*	0.07	*0.26*	0.07	*0.26*	0.06	*0.25*	0.06	*0.24*
	25-29	0.09	*0.28*	0.07	*0.25*	0.08	*0.28*	0.07	*0.25*	0.06	*0.24*
	30-39	0.20	*0.40*	0.14	*0.35*	0.16	*0.37*	0.19	*0.39*	0.14	*0.35*
	40-49	0.31	*0.46*	0.38	*0.48*	0.41	*0.49*	0.32	*0.47*	0.34	*0.47*
	50-59	0.34	*0.47*	0.30	*0.46*	0.20	*0.40*	0.31	*0.46*	0.34	*0.47*
	Over 60	0.04	*0.19*	0.04	*0.21*	0.06	*0.24*	0.04	*0.20*	0.06	*0.23*
	Female	0.80	*0.40*	0.78	*0.42*	0.66	*0.47*	0.78	*0.41*	0.64	*0.48*
Experience	1-2 years	0.06	*0.24*	0.11	*0.31*	0.11	*0.31*	0.09	*0.29*	0.10	*0.30*
	3-5 years	0.03	*0.18*	0.05	*0.22*	0.06	*0.24*	0.05	*0.21*	0.04	*0.20*
	Over 6 years	0.91	*0.29*	0.84	*0.37*	0.83	*0.38*	0.86	*0.35*	0.86	*0.35*
	University diploma	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.44	*0.74*	2.13	*0.87*	2.18	*0.90*	2.12	*0.83*	2.15	*0.81*
	Prepares material	2.52	*0.96*	2.49	*1.12*	2.57	*1.07*	2.38	*1.02*	2.50	*1.12*
	Visits other classes	1.48	*0.53*	1.48	*0.58*	1.46	*0.60*	1.39	*0.51*	1.45	*0.56*
	Informal visits	1.28	*0.46*	1.30	*0.46*	1.29	*0.45*	1.26	*0.45*	1.30	*0.52*
Further development	Subject content course	0.50	*0.50*	0.44	*0.50*	0.47	*0.50*	0.50	*0.50*	0.00	*0.00*
	Holds a certificate	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
	Curriculum improvement course	0.28	*0.45*	0.28	*0.45*	0.29	*0.45*	0.23	*0.42*	0.26	*0.44*
	Subject related to IT	0.27	*0.44*	0.36	*0.48*	0.33	*0.47*	0.29	*0.45*	0.29	*0.45*
	Critical thinking course	0.37	*0.48*	0.31	*0.46*	0.33	*0.47*	0.30	*0.46*	0.23	*0.42*
Class characteristics	Class size	20.09	*7.77*	19.00	*8.78*	18.55	*8.82*	19.01	*8.75*	18.78	*8.31*
	Minutes in class	164.43	*29.43*	67.45	*34.45*	62.11	*29.29*	62.85	*31.84*	60.24	*28.92*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.17: Descriptive statistics: Teacher and class characteristics in Indonesia

		Mathematics		Biology		Physics	
		146 Teachers		127 Teachers		129 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.14*	0.01	*0.11*	0.07	*0.25*
	25-29	0.13	*0.34*	0.18	*0.38*	0.16	*0.37*
	30-39	0.42	*0.49*	0.42	*0.49*	0.47	*0.50*
	40-49	0.37	*0.48*	0.34	*0.47*	0.24	*0.43*
	50-59	0.05	*0.22*	0.05	*0.21*	0.05	*0.21*
	Over 60	0.02	*0.13*	0.00	*0.00*	0.01	*0.12*
	Female	0.42	*0.49*	0.64	*0.48*	0.47	*0.50*
Experience	1-2 years	0.07	*0.25*	0.07	*0.26*	0.10	*0.30*
	3-5 years	0.13	*0.33*	0.17	*0.38*	0.12	*0.33*
	Over 6 years	0.81	*0.39*	0.75	*0.43*	0.78	*0.42*
	University diploma	0.89	*0.32*	0.89	*0.31*	0.83	*0.38*
Motivation on 1-4 scale	Discusses concepts with others	2.47	*0.89*	2.53	*0.78*	2.49	*0.81*
	Prepares material	3.04	*0.97*	3.01	*0.93*	2.80	*0.94*
	Visits other classes	1.64	*0.83*	1.71	*0.93*	1.53	*0.75*
	Informal visits	1.55	*0.74*	1.58	*0.79*	1.45	*0.68*
Further development	Subject content course	0.67	*0.47*	0.67	*0.47*	0.00	*0.00*
	Holds a certificate	0.87	*0.33*	0.89	*0.31*	0.87	*0.34*
	Curriculum improvement course	0.76	*0.43*	0.71	*0.45*	0.71	*0.46*
	Subject related to IT	0.29	*0.46*	0.28	*0.45*	0.24	*0.43*
	Critical thinking course	0.56	*0.50*	0.54	*0.50*	0.55	*0.50*
Class characteristics	Class size	37.04	*13.23*	35.20	*8.61*	34.15	*10.62*
	Minutes in class	202.62	*64.95*	123.46	*76.60*	126.22	*74.42*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.18: Descriptive statistics: Teacher and class characteristics in Islamic Rep. of Iran

		Mathematics		Science	
		191 Teachers		208 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.09	*0.29*	0.06	*0.24*
	25-29	0.20	*0.40*	0.10	*0.29*
	30-39	0.46	*0.50*	0.50	*0.50*
	40-49	0.18	*0.39*	0.24	*0.42*
	50-59	0.05	*0.21*	0.11	*0.31*
	Over 60	0.01	*0.09*	0.00	*0.00*
	Female	0.40	*0.49*	0.41	*0.49*
Experience	1-2 years	0.04	*0.20*	0.07	*0.25*
	3-5 years	0.16	*0.37*	0.10	*0.30*
	Over 6 years	0.80	*0.40*	0.83	*0.38*
	University diploma	1.00	*0.00*	1.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.22	*0.73*	2.32	*0.76*
	Prepares material	2.13	*0.89*	2.35	*0.82*
	Visits other classes	1.16	*0.42*	1.26	*0.58*
	Informal visits	1.29	*0.53*	1.34	*0.59*
Further development	Subject content course	0.75	*0.43*	0.80	*0.40*
	Holds a certificate	1.00	*0.00*	0.98	*0.13*
	Curriculum improvement course	0.46	*0.50*	0.59	*0.49*
	Subject related to IT	0.28	*0.45*	0.43	*0.49*
	Critical thinking course	0.48	*0.50*	0.53	*0.50*
Class characteristics	Class size	23.60	*7.39*	23.76	*7.64*
	Minutes in class	180.36	*52.88*	171.62	*59.53*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.19: Descriptive statistics: Teacher and class characteristics in Israel

		Mathematics		Science	
		243 Teachers		270 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.03	*0.18*	0.18	*0.38*
	25-29	0.10	*0.30*	0.09	*0.28*
	30-39	0.30	*0.46*	0.24	*0.43*
	40-49	0.34	*0.47*	0.30	*0.46*
	50-59	0.19	*0.39*	0.18	*0.39*
	Over 60	0.04	*0.20*	0.01	*0.10*
	Female	0.76	*0.43*	0.70	*0.46*
Experience	1-2 years	0.10	*0.30*	0.25	*0.43*
	3-5 years	0.09	*0.29*	0.06	*0.23*
	Over 6 years	0.81	*0.39*	0.70	*0.46*
	University diploma	0.95	*0.22*	0.82	*0.38*
Motivation on 1-4 scale	Discusses concepts with others	2.36	*0.77*	2.18	*0.83*
	Prepares material	2.42	*0.82*	2.14	*0.80*
	Visits other classes	1.16	*0.48*	1.09	*0.32*
	Informal visits	1.17	*0.47*	1.13	*0.38*
Further development	Subject content course	0.57	*0.49*	0.58	*0.49*
	Holds a certificate	0.98	*0.14*	0.83	*0.37*
	Curriculum improvement course	0.50	*0.50*	0.61	*0.49*
	Subject related to IT	0.33	*0.47*	0.45	*0.50*
	Critical thinking course	0.42	*0.49*	0.49	*0.50*
Class characteristics	Class size	32.79	*5.00*	23.74	*15.68*
	Minutes in class	218.60	*46.46*	126.39	*70.44*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.20: Descriptive statistics: Teacher and class characteristics in Japan

		Mathematics		Science	
		201 Teachers		165 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.06	*0.24*	0.03	*0.16*
	25-29	0.14	*0.35*	0.17	*0.38*
	30-39	0.29	*0.46*	0.25	*0.43*
	40-49	0.37	*0.48*	0.32	*0.47*
	50-59	0.11	*0.32*	0.23	*0.42*
	Over 60	0.02	*0.14*	0.01	*0.09*
	Female	0.42	*0.49*	0.15	*0.36*
Experience	1-2 years	0.13	*0.34*	0.08	*0.27*
	3-5 years	0.10	*0.30*	0.17	*0.38*
	Over 6 years	0.77	*0.42*	0.75	*0.43*
	University diploma	1.00	*0.05*	0.99	*0.08*
Motivation on 1-4 scale	Discusses concepts with others	2.16	*0.78*	1.93	*0.79*
	Prepares material	2.58	*1.02*	1.76	*0.84*
	Visits other classes	1.52	*0.79*	1.36	*0.65*
	Informal visits	1.48	*0.86*	1.24	*0.57*
Further development	Subject content course	0.76	*0.43*	0.64	*0.48*
	Holds a certificate	0.99	*0.07*	0.99	*0.10*
	Curriculum improvement course	0.30	*0.46*	0.33	*0.47*
	Subject related to IT	0.26	*0.44*	0.33	*0.47*
	Critical thinking course	0.39	*0.49*	0.14	*0.35*
Class characteristics	Class size	32.47	*8.19*	33.66	*6.85*
	Minutes in class	158.36	*25.03*	148.56	*22.47*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.21: Descriptive statistics: Teacher and class characteristics in Jordan

		Mathematics		Science	
		193 Teachers		200 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.13	*0.33*	0.09	*0.29*
	25-29	0.26	*0.44*	0.37	*0.48*
	30-39	0.37	*0.48*	0.36	*0.48*
	40-49	0.18	*0.38*	0.13	*0.33*
	50-59	0.06	*0.25*	0.05	*0.21*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.51	*0.50*	0.54	*0.50*
Experience	1-2 years	0.20	*0.40*	0.25	*0.43*
	3-5 years	0.18	*0.39*	0.28	*0.45*
	Over 6 years	0.61	*0.49*	0.47	*0.50*
	University diploma	0.87	*0.34*	0.96	*0.20*
Motivation on 1-4 scale	Discusses concepts with others	2.44	*0.83*	2.64	*0.85*
	Prepares material	2.34	*0.89*	2.73	*0.96*
	Visits other classes	1.91	*0.72*	1.88	*0.72*
	Informal visits	2.00	*1.06*	1.85	*0.87*
Further development	Subject content course	0.78	*0.41*	0.77	*0.42*
	Holds a certificate	0.80	*0.40*	0.72	*0.45*
	Curriculum improvement course	0.61	*0.49*	0.64	*0.48*
	Subject related to IT	0.64	*0.48*	0.58	*0.49*
	Critical thinking course	0.65	*0.48*	0.73	*0.45*
Class characteristics	Class size	32.26	*10.31*	31.39	*11.54*
	Minutes in class	224.29	*22.74*	224.39	*18.03*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.22: Descriptive statistics: Teacher and class characteristics in South Korea

		Mathematics		Science	
		227 Teachers		181 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.14*	0.06	*0.24*
	25-29	0.22	*0.41*	0.16	*0.36*
	30-39	0.30	*0.46*	0.26	*0.44*
	40-49	0.36	*0.48*	0.40	*0.49*
	50-59	0.09	*0.29*	0.11	*0.31*
	Over 60	0.01	*0.11*	0.01	*0.11*
	Female	0.61	*0.49*	0.62	*0.49*
Experience	1-2 years	0.14	*0.34*	0.17	*0.37*
	3-5 years	0.19	*0.39*	0.14	*0.35*
	Over 6 years	0.68	*0.47*	0.69	*0.46*
	University diploma	1.00	*0.00*	0.99	*0.11*
Motivation on 1-4 scale	Discusses concepts with others	2.06	*0.71*	2.34	*0.75*
	Prepares material	2.50	*0.83*	2.69	*0.80*
	Visits other classes	1.22	*0.45*	1.26	*0.47*
	Informal visits	1.14	*0.39*	1.18	*0.43*
Further development	Subject content course	0.50	*0.50*	0.49	*0.50*
	Holds a certificate	1.00	*0.07*	0.98	*0.12*
	Curriculum improvement course	0.41	*0.49*	0.34	*0.47*
	Subject related to IT	0.30	*0.46*	0.28	*0.45*
	Critical thinking course	0.21	*0.41*	0.38	*0.49*
Class characteristics	Class size	36.53	*5.13*	35.42	*8.10*
	Minutes in class	181.33	*13.80*	161.47	*43.32*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.23: Descriptive statistics: Teacher and class characteristics in Rep. of Kuwait

		Mathematics		Science	
		106 Teachers		158 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.01	*0.09*	0.23	*0.42*
	25-29	0.17	*0.38*	0.15	*0.35*
	30-39	0.51	*0.50*	0.33	*0.47*
	40-49	0.19	*0.39*	0.22	*0.41*
	50-59	0.11	*0.32*	0.05	*0.23*
	Over 60	0.00	*0.00*	0.03	*0.16*
	Female	0.52	*0.50*	0.39	*0.49*
Experience	1-2 years	0.13	*0.34*	0.41	*0.49*
	3-5 years	0.12	*0.33*	0.06	*0.24*
	Over 6 years	0.74	*0.44*	0.52	*0.50*
	University diploma	0.98	*0.14*	0.79	*0.41*
Motivation on 1-4 scale	Discusses concepts with others	2.96	*0.85*	2.69	*1.15*
	Prepares material	3.01	*0.91*	2.82	*1.19*
	Visits other classes	2.06	*0.65*	1.85	*0.72*
	Informal visits	1.57	*0.76*	1.66	*0.81*
Further development	Subject content course	0.59	*0.49*	0.43	*0.50*
	Holds a certificate	0.75	*0.43*	0.55	*0.50*
	Curriculum improvement course	0.27	*0.44*	0.25	*0.44*
	Subject related to IT	0.43	*0.50*	0.35	*0.48*
	Critical thinking course	0.65	*0.48*	0.36	*0.48*
Class characteristics	Class size	29.29	*5.27*	20.38	*16.66*
	Minutes in class	77.12	*74.40*	61.21	*68.73*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.24: Descriptive statistics: Teacher and class characteristics in Lebanon

		Mathematics		Biology		Chemistry		Physics	
		116 Teachers		164 Teachers		167 Teachers		164 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.04	*0.19*	0.17	*0.38*	0.14	*0.35*	0.18	*0.39*
	25-29	0.30	*0.46*	0.33	*0.47*	0.28	*0.45*	0.26	*0.44*
	30-39	0.27	*0.45*	0.28	*0.45*	0.34	*0.47*	0.31	*0.46*
	40-49	0.19	*0.39*	0.14	*0.34*	0.17	*0.38*	0.15	*0.35*
	50-59	0.20	*0.40*	0.05	*0.22*	0.05	*0.22*	0.07	*0.26*
	Over 60	0.00	*0.00*	0.03	*0.16*	0.01	*0.11*	0.02	*0.15*
	Female	0.42	*0.49*	0.69	*0.46*	0.63	*0.48*	0.47	*0.50*
Experience	1-2 years	0.19	*0.39*	0.21	*0.41*	0.21	*0.41*	0.23	*0.42*
	3-5 years	0.14	*0.35*	0.21	*0.41*	0.16	*0.37*	0.17	*0.38*
	Over 6 years	0.67	*0.47*	0.58	*0.49*	0.63	*0.48*	0.60	*0.49*
	University diploma	0.68	*0.47*	0.75	*0.43*	0.78	*0.41*	0.75	*0.43*
Motivation on 1-4 scale	Discusses concepts with others	2.63	*0.91*	2.44	*0.90*	2.35	*0.88*	2.43	*0.91*
	Prepares material	2.43	*1.13*	2.39	*1.02*	2.43	*1.05*	2.40	*1.08*
	Visits other classes	1.43	*0.69*	1.44	*0.71*	1.47	*0.72*	1.47	*0.74*
	Informal visits	1.69	*0.79*	1.60	*0.88*	1.61	*0.79*	1.54	*0.77*
Further development	Subject content course	0.63	*0.48*	0.63	*0.48*	0.54	*0.50*	0.00	*0.00*
	Holds a certificate	0.65	*0.48*	0.75	*0.43*	0.69	*0.46*	0.72	*0.45*
	Curriculum improvement course	0.52	*0.50*	0.51	*0.50*	0.46	*0.50*	0.45	*0.50*
	Subject related to IT	0.48	*0.50*	0.40	*0.49*	0.36	*0.48*	0.40	*0.49*
	Critical thinking course	0.64	*0.48*	0.56	*0.50*	0.50	*0.50*	0.53	*0.50*
Class characteristics	Class size	24.95	*9.15*	20.28	*12.63*	22.68	*14.31*	21.50	*11.62*
	Minutes in class	262.70	*53.99*	103.37	*95.36*	107.62	*99.19*	112.63	*95.14*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.25: Descriptive statistics: Teacher and class characteristics in Lithuania

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		205	Teachers	160	Teachers	163	Teachers	157	Teachers	152	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.15*	0.02	*0.15*	0.02	*0.16*	0.03	*0.16*	0.05	*0.23*
	25-29	0.06	*0.23*	0.05	*0.21*	0.07	*0.26*	0.03	*0.17*	0.04	*0.20*
	30-39	0.12	*0.32*	0.22	*0.41*	0.19	*0.39*	0.13	*0.33*	0.14	*0.35*
	40-49	0.46	*0.50*	0.35	*0.48*	0.41	*0.49*	0.42	*0.49*	0.39	*0.49*
	50-59	0.29	*0.46*	0.25	*0.43*	0.20	*0.40*	0.28	*0.45*	0.27	*0.44*
	Over 60	0.05	*0.21*	0.12	*0.32*	0.10	*0.30*	0.12	*0.32*	0.11	*0.31*
	Female	0.90	*0.29*	0.87	*0.34*	0.76	*0.43*	0.92	*0.27*	0.65	*0.48*
Experience	1-2 years	0.03	*0.18*	0.10	*0.30*	0.08	*0.28*	0.05	*0.22*	0.14	*0.35*
	3-5 years	0.04	*0.19*	0.05	*0.22*	0.07	*0.25*	0.06	*0.23*	0.04	*0.21*
	Over 6 years	0.93	*0.26*	0.85	*0.36*	0.85	*0.36*	0.89	*0.31*	0.81	*0.39*
	University diploma	0.93	*0.26*	0.92	*0.27*	0.93	*0.25*	0.92	*0.28*	0.89	*0.31*
Motivation on 1-4 scale	Discusses concepts with others	2.07	*0.81*	1.91	*0.60*	1.85	*0.67*	1.95	*0.70*	1.95	*0.75*
	Prepares material	2.13	*0.90*	2.19	*0.94*	2.07	*0.94*	2.14	*0.91*	2.15	*0.97*
	Visits other classes	1.38	*0.58*	1.46	*0.57*	1.38	*0.51*	1.45	*0.63*	1.45	*0.56*
	Informal visits	1.32	*0.57*	1.42	*0.56*	1.31	*0.46*	1.41	*0.56*	1.33	*0.56*
Further development	Subject content course	0.76	*0.43*	0.64	*0.48*	0.71	*0.45*	0.60	*0.49*	0.00	*0.00*
	Holds a certificate	0.96	*0.19*	0.98	*0.15*	0.95	*0.22*	0.96	*0.19*	0.95	*0.22*
	Curriculum improvement course	0.70	*0.46*	0.68	*0.47*	0.75	*0.43*	0.58	*0.49*	0.61	*0.49*
	Subject related to IT	0.70	*0.46*	0.66	*0.47*	0.60	*0.49*	0.71	*0.45*	0.68	*0.47*
	Critical thinking course	0.53	*0.50*	0.50	*0.50*	0.52	*0.50*	0.46	*0.50*	0.51	*0.50*
Class characteristics	Class size	23.06	*6.39*	21.38	*9.47*	21.64	*8.14*	21.97	*8.59*	22.10	*8.55*
	Minutes in class	182.14	*18.29*	46.69	*12.94*	90.03	*10.51*	89.26	*10.61*	89.17	*15.22*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.26: Descriptive statistics: Teacher and class characteristics in Malaysia

		Mathematics		Science	
		163 Teachers		163 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.14*	0.03	*0.16*
	25-29	0.21	*0.41*	0.20	*0.40*
	30-39	0.39	*0.49*	0.49	*0.50*
	40-49	0.28	*0.45*	0.22	*0.42*
	50-59	0.09	*0.29*	0.06	*0.23*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.71	*0.46*	0.75	*0.43*
Experience	1-2 years	0.10	*0.30*	0.08	*0.27*
	3-5 years	0.19	*0.39*	0.19	*0.39*
	Over 6 years	0.71	*0.45*	0.73	*0.44*
	University diploma	0.87	*0.33*	0.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.43	*0.65*	2.64	*0.74*
	Prepares material	2.27	*0.80*	2.64	*0.85*
	Visits other classes	1.33	*0.58*	1.45	*0.62*
	Informal visits	1.59	*0.64*	1.72	*0.63*
Further development	Subject content course	0.46	*0.50*	0.46	*0.50*
	Holds a certificate	0.99	*0.09*	0.97	*0.17*
	Curriculum improvement course	0.52	*0.50*	0.65	*0.48*
	Subject related to IT	0.62	*0.49*	0.60	*0.49*
	Critical thinking course	0.27	*0.45*	0.38	*0.49*
Class characteristics	Class size	35.23	*5.21*	34.59	*7.41*
	Minutes in class	189.38	*20.49*	183.34	*40.11*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.27: Descriptive statistics: Teacher and class characteristics in Malta

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		112	Teachers	72	Teachers	51	Teachers	49	Teachers	119	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.11	*0.32*	0.15	*0.36*	0.09	*0.29*	0.10	*0.30*	0.24	*0.43*
	25-29	0.34	*0.47*	0.23	*0.42*	0.42	*0.49*	0.14	*0.35*	0.29	*0.45*
	30-39	0.32	*0.47*	0.41	*0.49*	0.19	*0.39*	0.39	*0.49*	0.33	*0.47*
	40-49	0.14	*0.34*	0.14	*0.34*	0.21	*0.41*	0.20	*0.40*	0.10	*0.30*
	50-59	0.09	*0.28*	0.08	*0.26*	0.09	*0.28*	0.09	*0.29*	0.04	*0.20*
	Over 60	0.00	*0.00*	0.00	*0.05*	0.00	*0.00*	0.08	*0.27*	0.00	*0.07*
	Female	0.58	*0.49*	0.73	*0.44*	0.52	*0.50*	0.61	*0.49*	0.57	*0.50*
Experience	1-2 years	0.12	*0.33*	0.10	*0.30*	0.14	*0.35*	0.12	*0.32*	0.22	*0.41*
	3-5 years	0.20	*0.40*	0.16	*0.36*	0.19	*0.40*	0.15	*0.35*	0.17	*0.37*
	Over 6 years	0.68	*0.47*	0.75	*0.44*	0.67	*0.47*	0.74	*0.44*	0.61	*0.49*
	University diploma	0.92	*0.27*	0.93	*0.26*	0.98	*0.13*	0.97	*0.17*	0.86	*0.34*
Motivation on 1-4 scale	Discusses concepts with others	2.10	*0.81*	1.96	*0.93*	1.82	*0.70*	1.90	*0.91*	2.19	*0.81*
	Prepares material	1.93	*0.93*	1.95	*0.85*	1.58	*0.64*	1.76	*0.72*	2.02	*0.77*
	Visits other classes	1.03	*0.26*	1.09	*0.33*	1.04	*0.19*	1.09	*0.29*	1.05	*0.23*
	Informal visits	1.06	*0.32*	1.08	*0.27*	1.14	*0.43*	1.07	*0.26*	1.06	*0.24*
Further development	Subject content course	0.69	*0.46*	0.44	*0.50*	0.14	*0.35*	0.24	*0.43*	0.00	*0.00*
	Holds a certificate	0.94	*0.25*	0.89	*0.32*	0.97	*0.16*	0.86	*0.35*	0.79	*0.40*
	Curriculum improvement course	0.59	*0.49*	0.54	*0.50*	0.24	*0.43*	0.49	*0.50*	0.26	*0.44*
	Subject related to IT	0.83	*0.37*	0.34	*0.47*	0.38	*0.48*	0.27	*0.45*	0.33	*0.47*
	Critical thinking course	0.30	*0.46*	0.27	*0.44*	0.28	*0.45*	0.20	*0.40*	0.22	*0.42*
Class characteristics	Class size	21.68	*3.89*	14.31	*5.33*	18.77	*7.34*	12.73	*6.19*	17.71	*7.20*
	Minutes in class	213.27	*22.80*	160.96	*36.23*	60.01	*47.28*	154.03	*32.62*	159.59	*46.20*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.28: Descriptive statistics: Teacher and class characteristics in Mongolia

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		132	Teachers	38	Teachers	28	Teachers	39	Teachers	47	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.16	*0.36*	0.16	*0.37*	0.23	*0.42*	0.23	*0.42*	0.17	*0.38*
	25-29	0.29	*0.45*	0.17	*0.38*	0.11	*0.32*	0.15	*0.36*	0.23	*0.42*
	30-39	0.18	*0.38*	0.14	*0.34*	0.18	*0.38*	0.12	*0.33*	0.18	*0.38*
	40-49	0.22	*0.42*	0.40	*0.49*	0.30	*0.46*	0.41	*0.49*	0.31	*0.46*
	50-59	0.13	*0.34*	0.13	*0.33*	0.15	*0.35*	0.09	*0.28*	0.11	*0.32*
	Over 60	0.02	*0.14*	0.00	*0.00*	0.03	*0.16*	0.00	*0.00*	0.00	*0.00*
	Female	0.77	*0.42*	0.87	*0.34*	0.81	*0.39*	0.76	*0.43*	0.70	*0.46*
Experience	1-2 years	0.16	*0.36*	0.24	*0.43*	0.32	*0.47*	0.27	*0.44*	0.26	*0.44*
	3-5 years	0.16	*0.37*	0.09	*0.28*	0.11	*0.31*	0.05	*0.22*	0.08	*0.27*
	Over 6 years	0.68	*0.47*	0.67	*0.47*	0.57	*0.50*	0.68	*0.47*	0.66	*0.47*
	University diploma	0.94	*0.23*	0.00	*0.00*	0.80	*0.40*	0.00	*0.00*	0.80	*0.40*
Motivation on 1-4 scale	Discusses concepts with others	2.41	*0.73*	2.44	*0.73*	2.13	*0.80*	2.35	*0.88*	2.21	*0.84*
	Prepares material	2.19	*0.79*	2.32	*0.83*	2.37	*0.88*	2.47	*0.93*	2.17	*0.91*
	Visits other classes	2.10	*0.79*	2.22	*0.76*	2.18	*0.81*	2.21	*0.94*	1.94	*0.92*
	Informal visits	1.97	*0.81*	2.32	*0.81*	2.46	*1.12*	2.16	*1.04*	1.98	*0.96*
Further development	Subject content course	0.51	*0.50*	0.58	*0.49*	0.45	*0.50*	0.39	*0.49*	0.00	*0.00*
	Holds a certificate	0.85	*0.35*	0.83	*0.37*	0.82	*0.38*	0.76	*0.43*	0.79	*0.41*
	Curriculum improvement course	0.80	*0.40*	0.63	*0.48*	0.58	*0.49*	0.45	*0.50*	0.52	*0.50*
	Subject related to IT	0.37	*0.48*	0.24	*0.43*	0.22	*0.42*	0.17	*0.38*	0.31	*0.46*
	Critical thinking course	0.23	*0.42*	0.33	*0.47*	0.39	*0.49*	0.35	*0.48*	0.26	*0.44*
Class characteristics	Class size	31.66	*7.87*	28.21	*11.31*	25.62	*14.98*	27.87	*14.58*	26.00	*13.44*
	Minutes in class	173.62	*42.85*	103.37	*82.77*	91.50	*83.78*	117.69	*84.87*	100.95	*88.03*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.29: Descriptive statistics: Teacher and class characteristics in Oman

		Mathematics		Science	
		143 Teachers		149 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.20	*0.40*	0.22	*0.42*
	25-29	0.60	*0.49*	0.53	*0.50*
	30-39	0.14	*0.34*	0.17	*0.38*
	40-49	0.06	*0.24*	0.06	*0.24*
	50-59	0.00	*0.07*	0.01	*0.12*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.54	*0.50*	0.49	*0.50*
Experience	1-2 years	0.30	*0.46*	0.34	*0.47*
	3-5 years	0.26	*0.44*	0.29	*0.46*
	Over 6 years	0.44	*0.50*	0.37	*0.48*
	University diploma	0.98	*0.12*	0.94	*0.23*
Motivation on 1-4 scale	Discusses concepts with others	2.53	*0.80*	2.55	*0.82*
	Prepares material	2.29	*0.89*	2.80	*0.95*
	Visits other classes	1.81	*0.50*	1.80	*0.52*
	Informal visits	1.80	*0.80*	1.84	*0.89*
Further development	Subject content course	0.38	*0.48*	0.44	*0.50*
	Holds a certificate	0.90	*0.30*	0.84	*0.37*
	Curriculum improvement course	0.54	*0.50*	0.43	*0.50*
	Subject related to IT	0.22	*0.41*	0.20	*0.40*
	Critical thinking course	0.33	*0.47*	0.26	*0.44*
Class characteristics	Class size	29.62	*7.72*	28.79	*8.93*
	Minutes in class	244.87	*81.57*	199.01	*99.41*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.30: Descriptive statistics: Teacher and class characteristics in Palestinian Nat'l Auth.

		Mathematics		Science	
		143 Teachers		152 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.12	*0.32*	0.05	*0.22*
	25-29	0.27	*0.44*	0.32	*0.47*
	30-39	0.27	*0.45*	0.27	*0.45*
	40-49	0.23	*0.42*	0.24	*0.43*
	50-59	0.11	*0.31*	0.11	*0.31*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.50	*0.50*	0.53	*0.50*
Experience	1-2 years	0.23	*0.42*	0.24	*0.42*
	3-5 years	0.15	*0.35*	0.22	*0.42*
	Over 6 years	0.63	*0.48*	0.54	*0.50*
	University diploma	0.84	*0.37*	0.89	*0.32*
Motivation on 1-4 scale	Discusses concepts with others	2.45	*0.87*	2.56	*0.82*
	Prepares material	2.19	*0.85*	2.69	*0.97*
	Visits other classes	1.69	*0.57*	1.72	*0.62*
	Informal visits	1.74	*0.77*	1.80	*0.76*
Further development	Subject content course	0.48	*0.50*	0.48	*0.50*
	Holds a certificate	0.60	*0.49*	0.65	*0.48*
	Curriculum improvement course	0.36	*0.48*	0.36	*0.48*
	Subject related to IT	0.24	*0.43*	0.35	*0.48*
	Critical thinking course	0.44	*0.50*	0.47	*0.50*
Class characteristics	Class size	35.81	*8.74*	34.90	*9.96*
	Minutes in class	166.50	*70.55*	146.16	*70.77*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.31: Descriptive statistics: Teacher and class characteristics in Qatar

		Mathematics		Science	
		124 Teachers		132 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.01	*0.11*	0.11	*0.31*
	25-29	0.24	*0.43*	0.16	*0.37*
	30-39	0.39	*0.49*	0.45	*0.50*
	40-49	0.23	*0.42*	0.18	*0.39*
	50-59	0.13	*0.33*	0.08	*0.27*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.48	*0.50*	0.43	*0.50*
Experience	1-2 years	0.15	*0.35*	0.18	*0.38*
	3-5 years	0.14	*0.34*	0.15	*0.36*
	Over 6 years	0.72	*0.45*	0.68	*0.47*
	University diploma	0.97	*0.18*	0.89	*0.31*
Motivation on 1-4 scale	Discusses concepts with others	2.99	*0.90*	2.80	*1.00*
	Prepares material	2.58	*0.98*	2.82	*1.07*
	Visits other classes	1.90	*0.57*	1.85	*0.60*
	Informal visits	1.72	*0.81*	1.64	*0.64*
Further development	Subject content course	0.55	*0.50*	0.62	*0.48*
	Holds a certificate	0.87	*0.34*	0.75	*0.43*
	Curriculum improvement course	0.38	*0.49*	0.42	*0.49*
	Subject related to IT	0.53	*0.50*	0.51	*0.50*
	Critical thinking course	0.51	*0.50*	0.44	*0.50*
Class characteristics	Class size	25.99	*6.30*	21.04	*11.15*
	Minutes in class	217.75	*48.32*	135.43	*98.78*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.32: Descriptive statistics: Teacher and class characteristics in Romania

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		229 Teachers		179 Teachers		188 Teachers		170 Teachers		189 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.15*	0.07	*0.25*	0.04	*0.19*	0.05	*0.21*	0.02	*0.14*
	25-29	0.05	*0.22*	0.11	*0.32*	0.16	*0.36*	0.09	*0.29*	0.05	*0.21*
	30-39	0.22	*0.42*	0.21	*0.41*	0.32	*0.47*	0.16	*0.37*	0.14	*0.35*
	40-49	0.21	*0.41*	0.12	*0.33*	0.14	*0.34*	0.30	*0.46*	0.39	*0.49*
	50-59	0.36	*0.48*	0.42	*0.49*	0.28	*0.45*	0.33	*0.47*	0.34	*0.48*
	Over 60	0.13	*0.34*	0.06	*0.24*	0.07	*0.25*	0.07	*0.25*	0.06	*0.23*
	Female	0.58	*0.49*	0.84	*0.37*	0.57	*0.49*	0.77	*0.42*	0.63	*0.48*
Experience	1-2 years	0.08	*0.27*	0.18	*0.39*	0.14	*0.35*	0.14	*0.34*	0.07	*0.25*
	3-5 years	0.05	*0.21*	0.06	*0.23*	0.19	*0.39*	0.08	*0.26*	0.04	*0.19*
	Over 6 years	0.87	*0.33*	0.76	*0.43*	0.67	*0.47*	0.79	*0.41*	0.90	*0.30*
	University diploma	0.99	*0.12*	0.96	*0.21*	0.96	*0.20*	0.98	*0.13*	0.98	*0.14*
Motivation on 1-4 scale	Discusses concepts with others	2.54	*0.84*	2.52	*0.83*	2.56	*0.81*	2.60	*0.77*	2.55	*0.82*
	Prepares material	2.65	*0.97*	2.96	*0.98*	3.01	*0.92*	2.86	*0.94*	2.90	*0.90*
	Visits other classes	1.86	*0.72*	1.85	*0.66*	1.78	*0.60*	1.80	*0.60*	1.77	*0.55*
	Informal visits	2.14	*0.99*	2.30	*1.02*	2.24	*0.98*	2.38	*1.00*	2.38	*1.02*
Further development	Subject content course	0.51	*0.50*	0.46	*0.50*	0.54	*0.50*	0.52	*0.50*	0.00	*0.00*
	Holds a certificate	0.97	*0.16*	0.93	*0.26*	0.94	*0.24*	0.91	*0.29*	0.94	*0.24*
	Curriculum improvement course	0.51	*0.50*	0.37	*0.48*	0.47	*0.50*	0.44	*0.50*	0.46	*0.50*
	Subject related to IT	0.54	*0.50*	0.63	*0.48*	0.59	*0.49*	0.65	*0.48*	0.69	*0.46*
	Critical thinking course	0.54	*0.50*	0.42	*0.49*	0.53	*0.50*	0.50	*0.50*	0.52	*0.50*
Class characteristics	Class size	19.10	*5.70*	18.48	*6.50*	18.33	*7.22*	18.29	*7.34*	18.30	*7.19*
	Minutes in class	210.96	*41.42*	70.53	*45.14*	99.64	*35.24*	118.24	*52.06*	118.12	*51.79*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.33: Descriptive statistics: Teacher and class characteristics in Russian Federation

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		251	Teachers	223	Teachers	229	Teachers	226	Teachers	231	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.02	*0.12*	0.05	*0.23*	0.03	*0.18*	0.04	*0.19*	0.04	*0.19*
	25-29	0.03	*0.17*	0.04	*0.20*	0.08	*0.27*	0.03	*0.18*	0.09	*0.29*
	30-39	0.19	*0.40*	0.23	*0.42*	0.29	*0.45*	0.19	*0.39*	0.14	*0.35*
	40-49	0.35	*0.48*	0.30	*0.46*	0.30	*0.46*	0.34	*0.47*	0.34	*0.47*
	50-59	0.32	*0.47*	0.30	*0.46*	0.25	*0.43*	0.30	*0.46*	0.28	*0.45*
	Over 60	0.09	*0.28*	0.08	*0.27*	0.05	*0.22*	0.10	*0.30*	0.11	*0.31*
	Female	0.93	*0.26*	0.95	*0.22*	0.92	*0.27*	0.95	*0.21*	0.76	*0.43*
Experience	1-2 years	0.03	*0.16*	0.09	*0.29*	0.05	*0.23*	0.04	*0.19*	0.08	*0.27*
	3-5 years	0.03	*0.17*	0.03	*0.18*	0.03	*0.18*	0.04	*0.21*	0.05	*0.22*
	Over 6 years	0.94	*0.23*	0.88	*0.33*	0.91	*0.28*	0.92	*0.28*	0.87	*0.34*
	University diploma	0.99	*0.11*	0.97	*0.17*	0.96	*0.19*	0.96	*0.19*	0.97	*0.18*
Motivation on 1-4 scale	Discusses concepts with others	2.51	*0.85*	2.49	*0.78*	2.41	*0.77*	2.30	*0.67*	2.43	*0.80*
	Prepares material	2.43	*0.92*	2.52	*0.87*	2.52	*0.85*	2.49	*0.85*	2.45	*0.93*
	Visits other classes	1.99	*0.58*	2.03	*0.55*	2.07	*0.58*	1.98	*0.51*	1.99	*0.56*
	Informal visits	1.72	*0.61*	1.80	*0.60*	1.80	*0.65*	1.77	*0.60*	1.75	*0.63*
Further development	Subject content course	0.72	*0.45*	0.69	*0.46*	0.67	*0.47*	0.75	*0.43*	0.00	*0.00*
	Holds a certificate	1.00	*0.05*	0.93	*0.26*	0.95	*0.21*	0.91	*0.29*	0.95	*0.21*
	Curriculum improvement course	0.73	*0.45*	0.69	*0.46*	0.68	*0.47*	0.71	*0.45*	0.57	*0.49*
	Subject related to IT	0.64	*0.48*	0.63	*0.48*	0.60	*0.49*	0.65	*0.48*	0.63	*0.48*
	Critical thinking course	0.59	*0.49*	0.49	*0.50*	0.46	*0.50*	0.52	*0.50*	0.38	*0.49*
Class characteristics	Class size	19.31	*6.33*	19.41	*6.36*	19.39	*6.36*	19.41	*6.36*	19.33	*6.52*
	Minutes in class	230.41	*35.71*	86.30	*10.43*	86.11	*11.79*	94.82	*21.81*	87.82	*15.71*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.34: Descriptive statistics: Teacher and class characteristics in Saudi Arabia

		Mathematics		Science	
		148 Teachers		175 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.04	*0.19*	0.06	*0.23*
	25-29	0.32	*0.47*	0.21	*0.41*
	30-39	0.41	*0.49*	0.60	*0.49*
	40-49	0.15	*0.36*	0.12	*0.32*
	50-59	0.08	*0.26*	0.01	*0.11*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.45	*0.50*	0.47	*0.50*
Experience	1-2 years	0.19	*0.39*	0.24	*0.43*
	3-5 years	0.18	*0.39*	0.11	*0.31*
	Over 6 years	0.63	*0.48*	0.65	*0.48*
	University diploma	0.97	*0.18*	0.96	*0.20*
Motivation on 1-4 scale	Discusses concepts with others	2.29	*0.85*	2.30	*0.80*
	Prepares material	2.02	*0.95*	2.34	*0.98*
	Visits other classes	1.63	*0.65*	1.66	*0.64*
	Informal visits	1.57	*0.84*	1.52	*0.75*
Further development	Subject content course	0.45	*0.50*	0.54	*0.50*
	Holds a certificate	0.00	*0.00*	0.64	*0.48*
	Curriculum improvement course	0.18	*0.39*	0.21	*0.41*
	Subject related to IT	0.20	*0.40*	0.28	*0.45*
	Critical thinking course	0.33	*0.47*	0.43	*0.50*
Class characteristics	Class size	27.58	*13.10*	21.57	*15.85*
	Minutes in class	174.81	*44.50*	77.82	*86.86*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.35: Descriptive statistics: Teacher and class characteristics in Serbia

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		181	Teachers	177	Teachers	172	Teachers	162	Teachers	170	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.01	*0.10*	0.00	*0.00*	0.03	*0.17*	0.02	*0.14*	0.02	*0.13*
	25-29	0.07	*0.26*	0.03	*0.18*	0.09	*0.29*	0.04	*0.20*	0.02	*0.13*
	30-39	0.21	*0.41*	0.26	*0.44*	0.34	*0.47*	0.25	*0.43*	0.18	*0.38*
	40-49	0.20	*0.40*	0.29	*0.45*	0.21	*0.41*	0.19	*0.39*	0.31	*0.46*
	50-59	0.36	*0.48*	0.35	*0.48*	0.30	*0.46*	0.45	*0.50*	0.42	*0.49*
	Over 60	0.15	*0.36*	0.07	*0.25*	0.02	*0.15*	0.05	*0.21*	0.05	*0.23*
	Female	0.61	*0.49*	0.79	*0.41*	0.66	*0.48*	0.81	*0.39*	0.65	*0.48*
Experience	1-2 years	0.10	*0.29*	0.11	*0.32*	0.16	*0.37*	0.18	*0.38*	0.09	*0.29*
	3-5 years	0.10	*0.30*	0.11	*0.32*	0.14	*0.34*	0.10	*0.30*	0.07	*0.26*
	Over 6 years	0.81	*0.40*	0.78	*0.42*	0.70	*0.46*	0.72	*0.45*	0.83	*0.37*
	University diploma	0.98	*0.14*	0.00	*0.00*	0.96	*0.19*	0.99	*0.11*	0.98	*0.13*
Motivation on 1-4 scale	Discusses concepts with others	2.88	*0.92*	2.60	*0.92*	2.46	*0.89*	2.48	*0.85*	2.58	*0.83*
	Prepares material	2.32	*0.94*	2.41	*0.95*	2.30	*0.94*	2.38	*0.96*	2.29	*0.93*
	Visits other classes	1.24	*0.49*	1.30	*0.58*	1.29	*0.49*	1.25	*0.44*	1.20	*0.44*
	Informal visits	1.24	*0.44*	1.35	*0.63*	1.34	*0.49*	1.32	*0.52*	1.28	*0.56*
Further development	Subject content course	0.46	*0.50*	0.35	*0.48*	0.27	*0.45*	0.34	*0.47*	0.00	*0.00*
	Holds a certificate	0.84	*0.37*	0.94	*0.24*	0.82	*0.38*	0.84	*0.37*	0.86	*0.35*
	Curriculum improvement course	0.43	*0.50*	0.16	*0.36*	0.15	*0.36*	0.19	*0.39*	0.22	*0.41*
	Subject related to IT	0.34	*0.47*	0.43	*0.50*	0.48	*0.50*	0.51	*0.50*	0.50	*0.50*
	Critical thinking course	0.36	*0.48*	0.47	*0.50*	0.42	*0.49*	0.42	*0.49*	0.43	*0.49*
Class characteristics	Class size	22.33	*5.38*	21.20	*7.40*	20.60	*8.52*	20.75	*7.79*	20.75	*8.23*
	Minutes in class	180.61	*15.78*	89.37	*10.28*	85.98	*17.96*	87.14	*17.35*	86.84	*17.09*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.36: Descriptive statistics: Teacher and class characteristics in Scotland

		Mathematics		Science	
		279 Teachers		859 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.07	*0.26*	0.27	*0.44*
	25-29	0.14	*0.34*	0.13	*0.33*
	30-39	0.24	*0.43*	0.13	*0.34*
	40-49	0.25	*0.43*	0.20	*0.40*
	50-59	0.27	*0.44*	0.25	*0.43*
	Over 60	0.04	*0.20*	0.03	*0.17*
	Female	0.58	*0.49*	0.39	*0.49*
Experience	1-2 years	0.20	*0.40*	0.41	*0.49*
	3-5 years	0.11	*0.31*	0.07	*0.26*
	Over 6 years	0.69	*0.46*	0.51	*0.50*
	University diploma	0.98	*0.15*	0.76	*0.43*
Motivation on 1-4 scale	Discusses concepts with others	2.41	*0.86*	2.15	*1.00*
	Prepares material	2.48	*0.99*	2.10	*1.05*
	Visits other classes	1.37	*0.62*	1.33	*0.65*
	Informal visits	1.55	*0.72*	1.55	*0.84*
Further development	Subject content course	0.90	*0.30*	0.62	*0.48*
	Holds a certificate	1.00	*0.00*	0.77	*0.42*
	Curriculum improvement course	0.74	*0.44*	0.51	*0.50*
	Subject related to IT	0.78	*0.41*	0.50	*0.50*
	Critical thinking course	0.56	*0.50*	0.48	*0.50*
Class characteristics	Class size	22.16	*9.15*	12.94	*10.41*
	Minutes in class	212.75	*34.18*	112.60	*76.36*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.37: Descriptive statistics: Teacher and class characteristics in Singapore

		Mathematics		Science	
		321 Teachers		376 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.07	*0.26*	0.11	*0.31*
	25-29	0.39	*0.49*	0.35	*0.48*
	30-39	0.31	*0.46*	0.28	*0.45*
	40-49	0.11	*0.32*	0.13	*0.33*
	50-59	0.11	*0.31*	0.13	*0.33*
	Over 60	0.01	*0.10*	0.01	*0.11*
	Female	0.64	*0.48*	0.63	*0.48*
Experience	1-2 years	0.30	*0.46*	0.28	*0.45*
	3-5 years	0.33	*0.47*	0.27	*0.44*
	Over 6 years	0.37	*0.48*	0.45	*0.50*
	University diploma	0.96	*0.20*	0.91	*0.29*
Motivation on 1-4 scale	Discusses concepts with others	2.30	*0.72*	2.40	*0.80*
	Prepares material	2.33	*0.89*	2.40	*0.85*
	Visits other classes	1.24	*0.52*	1.33	*0.55*
	Informal visits	1.38	*0.55*	1.38	*0.57*
Further development	Subject content course	0.88	*0.32*	0.83	*0.37*
	Holds a certificate	0.98	*0.15*	0.96	*0.21*
	Curriculum improvement course	0.64	*0.48*	0.76	*0.43*
	Subject related to IT	0.74	*0.44*	0.68	*0.46*
	Critical thinking course	0.62	*0.49*	0.70	*0.46*
Class characteristics	Class size	37.94	*3.94*	37.10	*6.98*
	Minutes in class	217.80	*37.07*	184.94	*53.43*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.38: Descriptive statistics: Teacher and class characteristics in Slovenia

		Mathematics		Biology		Chemistry		Physics	
		443	Teachers	172	Teachers	157	Teachers	170	Teachers
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.01	*0.08*	0.04	*0.20*	0.03	*0.18*	0.02	*0.13*
	25-29	0.16	*0.37*	0.08	*0.26*	0.10	*0.29*	0.09	*0.29*
	30-39	0.24	*0.42*	0.25	*0.43*	0.24	*0.43*	0.25	*0.43*
	40-49	0.39	*0.49*	0.42	*0.49*	0.40	*0.49*	0.42	*0.49*
	50-59	0.20	*0.40*	0.21	*0.41*	0.23	*0.42*	0.18	*0.39*
	Over 60	0.01	*0.10*	0.00	*0.07*	0.00	*0.00*	0.03	*0.17*
	Female	0.83	*0.38*	0.88	*0.32*	0.89	*0.31*	0.63	*0.48*
Experience	1-2 years	0.09	*0.29*	0.11	*0.32*	0.08	*0.27*	0.07	*0.26*
	3-5 years	0.13	*0.34*	0.07	*0.26*	0.07	*0.26*	0.08	*0.27*
	Over 6 years	0.78	*0.42*	0.81	*0.39*	0.85	*0.36*	0.85	*0.36*
	University diploma	0.95	*0.23*	0.96	*0.19*	0.97	*0.16*	0.94	*0.24*
Motivation on 1-4 scale	Discusses concepts with others	2.60	*0.85*	2.22	*0.82*	2.13	*0.80*	2.18	*0.83*
	Prepares material	2.38	*0.88*	1.86	*0.81*	1.83	*0.79*	1.91	*0.84*
	Visits other classes	1.08	*0.34*	1.11	*0.35*	1.11	*0.36*	1.07	*0.26*
	Informal visits	1.18	*0.49*	1.17	*0.48*	1.19	*0.50*	1.12	*0.39*
Further development	Subject content course	0.65	*0.48*	0.58	*0.49*	0.47	*0.50*	0.00	*0.00*
	Holds a certificate	0.90	*0.30*	0.93	*0.25*	0.95	*0.21*	0.90	*0.30*
	Curriculum improvement course	0.64	*0.48*	0.49	*0.50*	0.37	*0.48*	0.36	*0.48*
	Subject related to IT	0.59	*0.49*	0.41	*0.49*	0.38	*0.49*	0.45	*0.50*
	Critical thinking course	0.35	*0.48*	0.28	*0.45*	0.18	*0.38*	0.24	*0.42*
Class characteristics	Class size	14.96	*5.66*	20.35	*5.88*	19.20	*7.34*	20.49	*5.66*
	Minutes in class	181.38	*12.50*	61.32	*36.30*	89.23	*29.04*	91.52	*21.67*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.39: Descriptive statistics: Teacher and class characteristics in Thailand

		Mathematics		Science	
		150 Teachers		150 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.05	*0.22*	0.04	*0.21*
	25-29	0.16	*0.37*	0.19	*0.39*
	30-39	0.29	*0.46*	0.30	*0.46*
	40-49	0.29	*0.46*	0.28	*0.45*
	50-59	0.20	*0.40*	0.18	*0.39*
	Over 60	0.00	*0.00*	0.00	*0.05*
	Female	0.64	*0.48*	0.65	*0.48*
Experience	1-2 years	0.17	*0.37*	0.12	*0.32*
	3-5 years	0.13	*0.33*	0.14	*0.35*
	Over 6 years	0.71	*0.46*	0.74	*0.44*
	University diploma	0.98	*0.14*	0.99	*0.10*
Motivation on 1-4 scale	Discusses concepts with others	2.61	*0.90*	2.55	*0.82*
	Prepares material	2.67	*0.82*	2.73	*0.84*
	Visits other classes	2.04	*0.76*	2.05	*0.83*
	Informal visits	1.68	*0.74*	1.66	*0.72*
Further development	Subject content course	0.74	*0.44*	0.80	*0.40*
	Holds a certificate	0.98	*0.15*	0.96	*0.19*
	Curriculum improvement course	0.76	*0.43*	0.78	*0.41*
	Subject related to IT	0.72	*0.45*	0.67	*0.47*
	Critical thinking course	0.80	*0.40*	0.78	*0.42*
Class characteristics	Class size	34.63	*10.01*	34.98	*10.10*
	Minutes in class	190.59	*56.15*	183.08	*47.22*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.40: Descriptive statistics: Teacher and class characteristics in Tunisia

		Mathematics		Science	
		139 Teachers		168 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.00	*0.00*	0.02	*0.14*
	25-29	0.18	*0.38*	0.18	*0.38*
	30-39	0.47	*0.50*	0.48	*0.50*
	40-49	0.21	*0.41*	0.23	*0.42*
	50-59	0.13	*0.34*	0.09	*0.29*
	Over 60	0.01	*0.09*	0.00	*0.00*
	Female	0.32	*0.47*	0.64	*0.48*
Experience	1-2 years	0.23	*0.42*	0.28	*0.45*
	3-5 years	0.14	*0.35*	0.20	*0.40*
	Over 6 years	0.64	*0.48*	0.53	*0.50*
	University diploma	0.95	*0.21*	0.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.38	*0.86*	2.58	*0.92*
	Prepares material	2.08	*0.91*	2.52	*0.93*
	Visits other classes	1.17	*0.59*	1.23	*0.50*
	Informal visits	1.50	*0.79*	1.40	*0.63*
Further development	Subject content course	0.32	*0.47*	0.75	*0.43*
	Holds a certificate	0.87	*0.34*	0.86	*0.35*
	Curriculum improvement course	0.20	*0.40*	0.72	*0.45*
	Subject related to IT	0.20	*0.40*	0.52	*0.50*
	Critical thinking course	0.37	*0.48*	0.39	*0.49*
Class characteristics	Class size	31.60	*4.82*	27.78	*11.52*
	Minutes in class	225.43	*32.91*	95.78	*36.59*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.41: Descriptive statistics: Teacher and class characteristics in Turkey

		Mathematics		Science	
		139 Teachers		146 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.18	*0.38*	0.07	*0.25*
	25-29	0.35	*0.48*	0.30	*0.46*
	30-39	0.14	*0.34*	0.35	*0.48*
	40-49	0.19	*0.39*	0.16	*0.37*
	50-59	0.15	*0.36*	0.13	*0.33*
	Over 60	0.00	*0.00*	0.00	*0.00*
	Female	0.49	*0.50*	0.50	*0.50*
Experience	1-2 years	0.27	*0.44*	0.22	*0.41*
	3-5 years	0.26	*0.44*	0.11	*0.32*
	Over 6 years	0.47	*0.50*	0.66	*0.47*
	University diploma	1.00	*0.00*	1.00	*0.00*
Motivation on 1-4 scale	Discusses concepts with others	2.40	*0.80*	2.40	*0.80*
	Prepares material	2.23	*0.82*	2.43	*0.80*
	Visits other classes	1.26	*0.49*	1.20	*0.49*
	Informal visits	1.17	*0.44*	1.14	*0.44*
Further development	Subject content course	0.48	*0.50*	0.64	*0.48*
	Holds a certificate	0.99	*0.10*	0.99	*0.09*
	Curriculum improvement course	0.67	*0.47*	0.77	*0.42*
	Subject related to IT	0.18	*0.38*	0.28	*0.45*
	Critical thinking course	0.23	*0.42*	0.25	*0.44*
Class characteristics	Class size	30.32	*9.84*	28.58	*10.91*
	Minutes in class	161.21	*9.32*	121.95	*16.63*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.42: Descriptive statistics: Teacher and class characteristics in Ukraine

		Mathematics		Biology		Earth Science		Chemistry		Physics	
		180 Teachers		184 Teachers		184 Teachers		184 Teachers		184 Teachers	
	Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	Under 25	0.04	*0.19*	0.06	*0.24*	0.08	*0.27*	0.07	*0.26*	0.04	*0.21*
	25-29	0.05	*0.22*	0.12	*0.32*	0.05	*0.22*	0.08	*0.27*	0.09	*0.29*
	30-39	0.23	*0.42*	0.27	*0.44*	0.27	*0.44*	0.24	*0.43*	0.25	*0.44*
	40-49	0.30	*0.46*	0.31	*0.46*	0.33	*0.47*	0.20	*0.40*	0.29	*0.45*
	50-59	0.30	*0.46*	0.18	*0.38*	0.24	*0.43*	0.31	*0.46*	0.20	*0.40*
	Over 60	0.08	*0.27*	0.07	*0.25*	0.03	*0.17*	0.10	*0.30*	0.12	*0.32*
	Female	0.91	*0.29*	0.90	*0.30*	0.83	*0.38*	0.94	*0.23*	0.61	*0.49*
Experience	1-2 years	0.04	*0.20*	0.09	*0.29*	0.08	*0.27*	0.08	*0.27*	0.03	*0.18*
	3-5 years	0.04	*0.20*	0.06	*0.23*	0.04	*0.20*	0.05	*0.21*	0.08	*0.27*
	Over 6 years	0.92	*0.28*	0.85	*0.35*	0.88	*0.33*	0.87	*0.33*	0.89	*0.32*
	University diploma	0.99	*0.09*	0.96	*0.20*	0.95	*0.21*	0.99	*0.11*	0.97	*0.17*
Motivation on 1-4 scale	Discusses concepts with others	2.82	*0.87*	2.71	*0.89*	2.58	*0.84*	2.74	*0.85*	2.75	*0.88*
	Prepares material	2.86	*0.92*	2.83	*0.95*	2.79	*0.98*	2.77	*0.92*	2.74	*0.94*
	Visits other classes	2.07	*0.49*	2.08	*0.49*	2.00	*0.39*	2.13	*0.47*	2.04	*0.38*
	Informal visits	1.81	*0.58*	1.78	*0.55*	1.79	*0.53*	1.87	*0.58*	1.80	*0.53*
Further development	Subject content course	0.79	*0.41*	0.82	*0.38*	0.79	*0.41*	0.81	*0.39*	0.00	*0.00*
	Holds a certificate	0.99	*0.12*	0.96	*0.19*	0.93	*0.25*	0.96	*0.21*	0.94	*0.23*
	Curriculum improvement course	0.81	*0.39*	0.82	*0.38*	0.80	*0.40*	0.83	*0.38*	0.82	*0.39*
	Subject related to IT	0.74	*0.44*	0.77	*0.42*	0.79	*0.41*	0.74	*0.44*	0.80	*0.40*
	Critical thinking course	0.79	*0.40*	0.73	*0.44*	0.71	*0.45*	0.82	*0.39*	0.81	*0.39*
Class characteristics	Class size	22.83	*7.25*	22.07	*8.14*	21.86	*8.60*	22.58	*7.63*	22.72	*7.57*
	Minutes in class	211.18	*37.41*	102.67	*40.52*	91.75	*31.52*	102.31	*37.40*	91.10	*13.19*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.

Table E.43: Descriptive statistics: Teacher and class characteristics in the United States

		Mathematics		Science	
		366 Teachers		497 Teachers	
	Variable	Mean	SD	Mean	SD
Age	Under 25	0.03	*0.18*	0.13	*0.34*
	25-29	0.16	*0.36*	0.11	*0.32*
	30-39	0.27	*0.44*	0.27	*0.44*
	40-49	0.26	*0.44*	0.23	*0.42*
	50-59	0.22	*0.41*	0.24	*0.43*
	Over 60	0.06	*0.23*	0.01	*0.12*
	Female	0.70	*0.46*	0.52	*0.50*
Experience	1-2 years	0.11	*0.32*	0.19	*0.39*
	3-5 years	0.14	*0.35*	0.17	*0.38*
	Over 6 years	0.74	*0.44*	0.64	*0.48*
	University diploma	1.00	*0.05*	0.90	*0.30*
Motivation on 1-4 scale	Discusses concepts with others	2.29	*0.97*	2.24	*1.01*
	Prepares material	2.14	*0.98*	2.25	*1.05*
	Visits other classes	1.25	*0.57*	1.31	*0.65*
	Informal visits	1.27	*0.60*	1.29	*0.61*
Further development	Subject content course	0.71	*0.45*	0.58	*0.49*
	Holds a certificate	0.96	*0.21*	0.85	*0.36*
	Curriculum improvement course	0.76	*0.43*	0.73	*0.45*
	Subject related to IT	0.57	*0.50*	0.64	*0.48*
	Critical thinking course	0.62	*0.48*	0.65	*0.48*
Class characteristics	Class size	21.41	*8.04*	21.69	*15.77*
	Minutes in class	246.04	*80.16*	196.55	*99.69*

Note: Teacher weights were used to calculate descriptive statistics. All variables have values 0 or 1, except for the following. Class characteristics are numerical values. Variables motivation range from values 1 to 4 where 1 corresponds to 'never or almost never', 2 corresponds to 'two or three times per month', 3 correspods to 'one up to three times a week', 4 corresponds to 'daily or almost daily'). Standard deviation (SD) in italics.