Institute of Economic Studies, Faculty of Social Sciences
Charles University in Prague

Do Teaching Practices Impact Socio-emotional Skills?

Vaclav Korbel
Michal Paulus

IES Working Paper: 4/2017

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Korbel V., Paulus M. (2017). "Do Teaching Practices Impact Socio-emotional Skills?" IES Working Paper 4/2017. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

Do Teaching Practices Impact Socio-emotional Skills?

Vaclav Korbel[a,b]
Michal Paulus[a,c]

[a]Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Smetanovo nabrezi 6, 111 01 Prague 1, Czech Republic
Email (corresponding author): vaclav.korbel@fsv.cuni.cz

February 2017

Abstract:
Recent studies emphasize the importance of socio-emotional skills for the success in school as well as for later economic outcomes. However, little is known how practices used by teachers everyday in classrooms impact socio-emotional skills. We show that modern practices such as working in small groups improve them. Especially intrinsic motivation and self-confidence are positively affected. Moreover, modern practices have no adverse effects on test scores. Standard practices such as lecturing or memorizing have no impact on socio-emotional skills. Splitting the sample reveals detrimental effects of standard practices on socio-emotional outcomes of boys and positive for high-achieving girls. On the contrary, both genders gain similarly from modern practices.

Keywords: Teaching practices, socio-emotional skills, between-subject variation
JEL: I21, C23

Acknowledgements: We want to thank Aline Bütikofer, Barbara Pertold-Gebicka, Daniel Münich, Katrine Vellesen Løken and an anonymous referee for valuable comments. All errors remaining in this text are our own. The research was supported by GA UK grant No. 910794.

1 Introduction

Socio-emotional skills are important determinants of school and later economic outcomes. They actuate children's grades, probability of high-school graduation or their future earnings (Cunha and Heckman, 2006). Several studies also supported their predictive ability towards risky behavior (Heckman et al., 2006). Crucially, socio-emotional skills are malleable primarily during childhood and adolescence (for a review see Heckman et al. (2010)) and can therefore be influenced in the educational process (Koch et al., 2015). However, little is known how teachers and the way they teach every day in classrooms affect formation of these skills.

Psychology examines the importance of various socio-emotional skills for schooling for decades while economic literature turned to it quite recently. Wigfield et al. (2009) show that motivation improves engagement in learning activities, academic performance or graduation rates. Similarly, self-control is not only a predictor of academic success (Mischel et al., 1989, Tangney et al., 2004) but also of health or public safety (Golsteyn et al., 2014, Moffitt et al., 2011). Borghans et al. (2008) look at personality traits where they find, using a taxonomy of the Big Five, that conscientiousness is a better predictor of educational attainment than cognitive abilities. Performance and beliefs of students can even be influenced by more subtle changes in, for example, self-confidence or self-concept (Meece et al., 1990, Spencer et al., 1999). Improvement in socio-emotional skills can even have wider effects across population than sole focus on content knowledge or cognitive abilities such as IQ which is more relevant for occupations with higher complexity (Kautz et al., 2014).

Research has shown that good teachers produce well performing students but so far it is unable to identify what makes a teacher good. After failing to find a systematic relationship of teacher characteristics with performance of students (Rockoff, 2004, Staiger and Rockoff, 2010) the focus was shifted to what teachers actually do in classes - to teaching practices. Researchers were looking primarily on how impacts of standard (traditional) practices like lecturing in front of the classroom or memorizing facts and formulas differ from modern practices like group

projects or making the content more applicable for the real life. The evidence on achievement and teaching practices suggests that students benefit more from standard practices but it is far from conclusive. Schwerdt and Wuppermann (2011) and Bietenbeck (2014) find a positive effect of standard teaching practices on student performance. Lavy (2015) shows a positive effect of standard teaching practices on students from low socioeconomic background but positive effects of modern teaching practices on students from educated families. Klaveren (2011) finds no relationship between teaching practices and student performance.

However, it could be argued that test scores may capture skills taught by standard practices like content knowledge whereas modern teaching practices may promote other skills. Bietenbeck (2014) supports it by showing that modern teaching practices improve reasoning skills and application of knowledge (fluid intelligence). Algan et al. (2013) find in cross-country evidence that modern teaching practices positively influence social capital.

This paper tries to fill the gap in the existing literature examining the impact of teaching practices on socio-emotional skills. Our focus is specifically on intrinsic motivation, extrinsic motivation and self-confidence. We hypothesize that modern teaching methods promote these skills. To look for potentially negative effects of modern teaching practices we examine their effects on test scores. In case of negative effects policies supporting modern practices would be prone to criticism.

To test our hypotheses, we use data from 2007 wave of the Trends in International Mathematics and Science Study (TIMSS) for the Czech Republic. The dataset contains test scores and self-reported answers on motivation and self-confidence from five subjects (math, physics, biology, chemistry, earth science). Furthermore, data on teaching practices from student questionnaires allows us to construct class aggregated indices for standard and modern teaching practices. The index can be interpreted as an effective share of lesson taught by standard, modern and other teaching practices in each subject. Our empirical strategy relies on within-student between-subject variation which controls for most of the selection effects. Including

rich set of teacher characteristics and class variables further limits the problem that effects are driven by unobserved teacher characteristics.

In the last part, we extend our analysis for seven Central and Eastern European (CEE) countries. Importantly for our empirical analysis, all CEE countries have data for all five subjects.[1] Running the analysis on wider sample of countries would be problematic since in majority of countries students are observed only in two subjects (math, science). Moreover, CEE countries have common recent history (part of communist area until 1989) and similar development in education systems after the fall of iron curtain (Fiszbein, 2001). This fact decreases differences due to cultural or institutional specifics. CEE countries, similarly to the Czech Republic, are at the bottom in terms of motivation of students in learning among OECD countries (Mullis et al., 2012b, OECD, 2016).

We find that modern teaching practices have significant and sizeable impacts on socio-emotional skills, especially on intrinsic motivation and self-confidence. 10 percentage point rise (6 minutes in 60 minutes lesson) in modern teaching practices increases intrinsic motivation by 0.24 of standard deviation and confidence by 0.16 of standard deviation. Neither standard nor modern teaching practices affect test scores. This evidence supports our initial hypothesis that modern teaching practices improve socio-emotional skills without harming test scores. Interestingly, standard practices reduce socio-emotional outcomes for boys and boost for girls. The effect is primarily driven by high-achieving girls. External validity of our results is corroborated by analysis of Central and Eastern European countries where we find qualitatively similar results.

Our paper contributes to the existing literature in three ways. First, to our best knowledge it is the first paper analysing the relationship between teaching practices and socio-emotional skills. The only similar papers looking at different outcomes than test scores are Bietenbeck (2014) who exploits various domains of cognitive skills and Algan et al. (2013) who look at social capital. Second, our results

[1] Except for Slovenia which misses the data on Earth science.

suggest that because of more widespread usage of standard practices girls (mainly high-achieving) get an advantage in socio-emotional skills which may contribute to gender differences in education (Cornwell et al., 2013, Jacob, 2002). Increase in modern practices might help both girls and boys in development of socio-emotional skills. Third, our results support the evidence that test scores do not capture all skills developed in school.

The remainder of the paper is structured as follows. Section 2 discusses dataset and descriptive statistics. Section 3 describes empirical strategy. Results are presented in Section 4 together with cross-country analysis and robustness checks. Section 5 concludes.

2 Data

We use a representative sample of the Czech students from the 2007 wave of TIMSS testing.[2] It is the last wave where students reported teaching practices. The test was conducted with fourth- and eight-grade students. We limit our analysis on eight-graders because fourth-graders in the Czech Republic are typically taught by a single teacher in all subjects where within-student between-subject approach could not be employed.

Students were tested in five subjects - mathematics, physics, biology, chemistry and earth science. The key element of the dataset is the possibility to link students to teachers and their practices in each subject.[3] TIMSS collects data by two-stage cluster sampling design. First, schools are chosen and then one or two classes are randomly drawn within school. Therefore, sampling weights and re-sampling techniques for variance estimation are used throughout the whole analysis.

The impact of teaching practices is measured on motivation, self-confidence and test scores. Motivation and self-confidence are derived from student self-reported values ranging from strongly agree (4) to strongly disagree (1). All variables includ-

[2]In total, 59 countries participated in the 2007 wave.

[3]Not all students are taught by five different teachers, 44% of students are taught by 4 different teachers and 13% by three or two teachers.

ing test scores are standardized to have mean 0 and standard deviation equal to 1.[4] We further divide motivation into intrinsic and extrinsic. Intrinsic motivation can be defined as a curiosity and joy of learning. In our analysis we use the question "I enjoy learning *subject*". Extrinsic motivation refers to a situation when a student is motivated to learn a subject because of an external goal. Economic literature studies extrinsic incentives mostly in the form of monetary rewards or grades (Dubey and Geanakoplos, 2010, Gneezy et al., 2011) but we focus on future prospects in education and job career. It is expressed by two questions: "I need to do well in *subject* to get into the university of my choice" and "I need to do well in *subject* to get the job I want". Self-confidence is derived from the question "*Subject* is more difficult for me than for many of my classmates" which is rescaled so that higher value means increase in self-confidence.[5]

Student questionnaires contain questions on teaching practices examining how often is a given practice used in lesson. We classify three teaching practices as standard (We listen to the teacher giving a lecture-style presentation, We memorize formulas and procedures, We work out problems on our own) and three as modern (We explain our answers, We relate what we are learning in *subject* to our daily lives, We work together in small groups/We work in small groups on an experiment or investigation).[6] Answers were in the first step rescaled so that "never" is equal to 0, "some lessons" to 0.25, "about half the lessons" to 0.5, and "every or almost every lesson" to 1. Rescaled variables represent effective share of lesson taught by standard and modern teaching practices on the individual level. Following Bietenbeck (2014), we aggregate the indices on a class level, leaving out the value for the observed pupil for each observation.[7]

[4]Section 4.3 compares linear estimation with a range of alternative fixed-effects ordered logit model estimators. Those models use original four point scale.

[5]Results are qualitatively and quantitatively similar when using alternative definitions of intrinsic motivation ("I find *subject* boring", "I like *subject*") and self-confidence ("*Subject* is not my strength").

[6]Modern teaching practices in math and science subjects slightly differ from each other - "We work together in small groups" in math and "We work in small groups on an experiment or investigation" in science subjects. However, both practices represent the same activity, namely, working in a small group.

[7]Results are robust when indices are created from simple means.

Table 1 summarizes standard and modern teaching practices across subjects. Standard practices are more widespread among Czech teachers. On average, teachers spend 64% of lesson teaching with standard practices and 44% with modern practices. Only 1.2% of teachers use modern practices more often than standard. Other activities not fitting into either category (controlling or doing homework, writing test or quizzes and using computers) take up on average 28% of lesson. It is important to note that sum of all activities does not have to add up one. Imagine a situation when a teacher relates the content to the real life (modern teaching practice) while giving a lecture (standard teaching practice) or when a student explains an answer (modern teaching practice) while reviewing homework (other activity). Our claim that standard and modern practices do not crowd-out each other is supported by their positive correlation reaching 0.45.

Final dataset consists of 22,633 observations representing 4,528 students in 212 classes taught by 711 teachers. Not to further decrease number of observations, missing values in all control variables are imputed with 0 and indicators for imputed values are used in all regressions.[8] TIMSS questionnaires contain a rich set of teacher characteristics including their motivation and further development. Class characteristics are matched from school questionnaires. Means and standard deviations are reported in Table 2. Age composition and share of female teachers significantly differ across subjects. Biology teachers participate the most in further development courses and earth science teachers the less. Importantly, teaching time is significantly higher for mathematics which may influence both motivation and test scores of students (Joyce et al., 2015). To control for confounding factors, we include variables presented in Table 2 in all regressions.

[8]The original dataset of 24,225 observations is restricted for missing values in all outcome variables and teaching practices containing 93 % of observations. Results are robust for dropping all missing values.

3 Empirical strategy

Natural way to estimate the impact of teaching practices on other outcomes would be to use standard education production function and regress variables of our interest on school characteristics, teacher characteristics and student characteristics. This approach would however neglect selection problems common in schools. First, if teaching practices are determined based on an unobserved school rule or teacher characteristic, then our estimates would be biased. For example, teachers who prefer modern teaching practices can self-select to schools with emphases on modern style of teaching. In the same vein, teachers preferring modern practices can be assigned within school to more motivated or able classes. Second, students may choose schools or classes based on style of teaching used there.

To avoid problems with selection, we use within-student between-subject approach following Aslam and Kingdon (2011). Comparing motivation or test scores across subjects eliminates any difference by between school or between class differences.[9] Nevertheless, we have to make an assumption that school and student characteristics influence our outcome variables similarly across subjects. Our student fixed-effects model can be written in the following way:

$$Y_{ijt} = \alpha_j + \beta T_{jt} + \gamma_1 STP_{ijt} + \gamma_2 MTP_{ijt} + \delta_i + \mu_{ijs}. \quad (1)$$

The outcome variable (motivation, self-confidence or test score) Y_{ijt} of student i in subject j taught by a teacher t is regressed on a vector of teacher characteristics T_{jt} and standard and modern teaching practices (STP_{ijt}, MTP_{ijt}). δ_i stands for the student fixed-effects and μ_{ijt} is the error term.

The within-student between-subject approach has two caveats. First, there could still be unobserved teacher characteristics influencing both outcome variable and selection of teaching practices. For example, more motivated or able teach-

[9] In our sample, almost all classes (99%) remain the same across subjects. The fact that Czech children are exposed in all classes to the same peers alleviates the chance that students benefit differently when exposed to better or worse peers (Hoxby, 2000).

ers may use modern practices more frequently. If unobserved motivation or skills promote outcome variables of student in another way, then our estimates would be biased. In the analysis, we control for a rich set of teacher characteristics including proxies for motivation and effort which should eliminate the problem. However, we cannot completely rule out a potential bias.

Second concern stems from potentially subject-specific selection of students to teaching practices. In other words, students may sort into schools that put emphases on a certain type of teaching practices in some of their subjects. Even though subject-specific sorting is not common in the Czech Republic,[10] some schools focus on specific subjects (e.g. languages or math). This could potentially be related to the choice of teaching practices. Section 4.3 investigates robustness of results for classes with and without special subject focus.

4 Results

We estimate the effects of standard (SP) and modern teaching practices (MP) on several outcome variables: test scores, intrinsic motivation, two types of extrinsic motivation and self-confidence. Results from our main specification are summarized in Table 3. Both teaching practice indices are included in all regressions.[11] All estimations are based on the student fixed-effect model and adjusted for the complex sampling design.

The estimated coefficients of teaching practices are interpreted in the following way. A X percentage point increase in a teaching variable (time spent on selected practices) increases outcome variable by $X * \beta$ of a standard deviation. For exam-

[10] In grade 6 and 8, students sort into grammar schools. Grammar schools are focused on preparing students to enter a university. Students (approximately 8% of our sample) sort there typically based on their skills and motivation through admission exams. However, the selection is not subject-specific.

[11] It could be argued that modern and standard practices are substitutes and so increase in one decreases the other. Even though positive correlation between both practices (0.45) suggests the opposite, we estimated equation 1 separately for each variable. Results are reported in Table A4, Panel A and B. Coefficients of modern practice index remain almost identical but coefficients of standard practice index increase for all socio-emotional skills. It indicates that both variables are interrelated and omitting one from the estimated equation would lead to bias in coefficients.

ple, 10 percentage point increase in modern practices accounts for 0.24 increase of standard deviation in intrinsic motivation. 10 percentage point change can be also expressed as 6 minutes in 60 minutes lesson. The resulting effect holds when standard practices are constant and vice versa. The change in time spent on modern or standard practices are at the expense of other teaching practices such as writing tests, reviewing homeworks or classroom management.

In general, the results reveal strong positive impact of modern practices on socio-emotional skills (Table 3). The highest impact is found for intrinsic motivation of the size 0.24 of SD (column 2). For extrinsic motivation (column 3-4), we observe positive influence of both practices, regardless if extrinsic motivation is related to future studies or job. In case of the first type of extrinsic motivation, the hypothesis that SP and MP coefficients are equal cannot be rejected. Their effect is about 0.04 SD when teaching practices increase by 10 percentage points (p-value=0.988). Impact on the students' motivation to get to the desired university is slightly higher than in case of their future job prospects (effect of 0.07 SD compared to 0.04 SD). Interestingly, additional time devoted to modern teaching practices strengthens self-confidence of students while standard practices have the opposite effect (column 5).[12]

Except for the effects of modern practices on socio-emotional skills we are also interested in their influence on test scores to check for potential adverse effects. Our estimation confirms no negative effects of MP on test-scores because coefficients of both practices are insignificant (column 1).[13] Our findings are in line with Klaveren (2011) who found no statistically significant relationship for students in Netherlands

[12]Table A3 estimates heterogeneity in effects across subjects. Math is the only subject where gains from standard practices are significantly higher and from modern practices significantly lower for motivation than in other subjects. Coefficients of other subjects are in accordance with previous findings. It suggests that modern practices are especially important for socio-emotional skills in science subjects.

[13]Table A2 documents relationship between teaching practices, socio-emotional skills and test scores. Not surprisingly, higher test scores are associated with higher motivation and self-confidence. The highest coefficient is for intrinsic motivation. In a similar vein, higher socio-emotional skills are associated with higher test scores. The size of coefficients for teaching practices remain almost identical in terms of socio-emotional skills. It seems implausible that effects of teaching practices are driven via effects on test scores. However, it is impossible disentangle the relationship among aforementioned variables with the dataset at hand.

and in contradiction with Schwerdt and Wuppermann (2011) or Bietenbeck (2014) who estimated positive effects of standard practices for US students. Positive effects of both practices on students in Israel was found by Lavy (2015). Mixed evidence hinders generalisation for policy makers and more international evidence is needed to explore what causes differences across countries.[14,15]

Interestingly, teacher characteristics have no or very little impact on test scores and socio-emotional skills. There is a positive effect of first few years of experience on intrinsic motivation but not on other outcomes. Surprisingly, teachers with university diploma motivate students less than those without a diploma. However, most of teachers (95%) have a diploma and the effect could be driven by a few teachers without a diploma who can compensate for it by other skills such as motivation or effort. Teaching practices turn out to be the most important variable for all socio-emotional outcomes.

4.1 Gender differences

Girls outperform boys in many countries both in math and science (Mullis et al., 2012a,b). Moreover, boys tend to be more often low achievers, more often drop out of school and have more behavioral problems (Bertrand and Pan, 2013, OECD, 2015). Even though the literature has mostly concentrated on reasons why girls could not catch up with boys in the past and still do not pursuit higher education more often (Beilock et al., 2010, Niederle and Vesterlund, 2010, Spencer et al., 1999), recent literature also explores why boys are starting to fall behind. Some authors argue that differences in socio-emotional skills explain a large share of the

[14]Table A4, Panel C investigates non-linearities in the relationship of outcome variables and teaching practices. Squared terms of teaching practices in the regressions turned out to be insignificant. We therefore cannot claim that the relationship is concave. Similar results were obtain also with other functional forms. Results available upon request.

[15]In some cases, we can observe teachers in two classes. Then, we can compare teaching practice indices between classes. In total, 156 teachers (22 %) in the dataset teach the same subject in two different classes. Correlation between indices is high reaching 0.57 for standard practices, 0.63 for modern practices and 0.67 for other practices. It suggests that our indices are adequate measures of teaching practices. Unfortunately, observing one or two teachers out of five subjects does not allow us to estimate coefficients with indices from the other class as suggested by Kane et al. (2011).

gap (Cornwell et al., 2013, Jacob, 2002). Education literature supports this claim showing that problems of boys are correlated with low intrinsic motivation and disinterest in school (Gorard et al., 1999, Houtte, 2004).

To explore aforementioned questions, we split the dataset by gender and performance (Table 4). Modern teaching practices have significant and positive effects for both genders but girls benefit significantly more from standard practices than boys (Panel A, B). Boys are in case of intrinsic motivation and self-confidence even harmed by usage of standard practices (Panel B). For girls, coefficients of standard practice indices are smaller or the same as coefficients of modern practice indices. When we look at girls divided by the mean test score, the effects are driven solely by high-achieving girls. When we divide boys in the same way we see no difference. It suggests that high-achieving girls receive socio-emotional "boost" from standard practices in science subjects and math. Therefore, increase in modern teaching practices may help boys without hurting girls. On the other hand, this result may be specific for science and math and does not have to hold in other subjects.

4.2 Central and Eastern Europe

In this section, we look at generalizability of our results for the region of Central and Eastern Europe. We include Bulgaria, Hungary, Slovenia, Lithuania, Romania, Russia, Ukraine and pooled sample of all CEE countries except for the Czech Republic. Those are all CEE countries participating in 2007 wave of TIMSS testing. Students from all countries are observed for all five subjects.[16]

Regressions estimate the equation 1 using the same set of control variables as for the Czech Republic. Results (Table 6) support our previous findings. Modern practices have positive impacts on all socio-emotional skills and even on test scores (Pooled data) which are significantly higher than for standard practices. The only exception is extrinsic motivation (university) where the difference is insignificant. Looking at individual countries, only Bulgarian students benefit more from stan-

[16]Except for Slovenia which misses data for Earth Science.

dard practices and Lithuanian and Romanian students benefit similarly from both practices. In other countries, null hypothesis of the same effect of two treatment variables is rejected. In conclusion, implications for promotion of modern teaching practices can be generalized for the region of CEE. However, national specifics need to be taken into account, especially in the case of Bulgaria.

4.3 Robustness checks

This subsection tests sensitivity of our results: first, to alternative definitions of teaching practices and second, to self-selection of students to teaching practices in the subject-specific way. All results are reported in Table 7. In the third part, we compare results from linear and ordered-logit models for socio-emotional outcomes (Table 8).

Impacts of standard and modern teaching practices on other outcomes may hinge on the exact composition of the teaching indices. Since students are asked on 16 practices, we further explore the robustness of our results by estimating our main specifications with two alternative teaching practice indices. They are presented in Table A1. The first one adds one standard and one modern practice. Second, teaching practice "work alone" could potentially be considered as both standard and modern, depending on the context, therefore we replace it with another practice. Panels A1 and A2 of Table 7 corroborate our original results with no effects on test scores and positive effects of modern practices on socio-emotional skills.

Within-student between-subject approach relies on the assumption that students do not sort to teaching practices in a subject-specific way. This would be violated if students would sort into schools or classes focusing on a specific subject/s. Moreover, emphases on a subject/s would have to be related to teaching practices. As we mentioned in the section 3, such kind of sorting is not common in the Czech Republic. Specifically, 8% classes in our sample specialize on a technical subject (math, ICT, or one of science subjects). When we divide classes on specialized and non-specialized, the results are qualitatively similar but the lower number

of observations leads to significance of fewer variables (Table 7, panel B1). The only important distinction is a positive effect of modern practices on test scores.[17] Similarly, the effects are not driven by grammar schools (selective schools for high achieving students) where the selection is not subject specific (Panel B2).

4.3.1 Linear and ordered-logit fixed-effects models

For our analysis, we assumed cardinality of answers in variables for socio-emotional skills. However, economists usually assume that self-reported answers are only ordinarily comparable. So far, evidence is not conclusive on what is a better approach. Some studies, exploring mainly happiness measures, argue that both approaches yield similar results (Dickerson et al., 2014, Ferrer-i Carbonell and Frijters, 2004, Frey and Stutzer, 2000). However, Baetschmann et al. (2015) showed in a recent study that fixed-effects (FE) ordered model used in Ferrer-i Carbonell and Frijters (2004) is biased. Then a comparison with cardinal results is not reliable.

In TIMSS questionnaires, answers range from "agree a lot" to "disagree a lot" which is strictly speaking rather an ordinal measure. On the other hand, FE ordered models have one important disadvantage that coefficients cannot be interpreted except for their sign and significance level. Because of simplicity of interpretation, we chose FE linear approach for our analysis. In this section we reanalyse all results with ordered FE approach. FE linear model is compared with four FE ordered-logit models - Blow-up and Cluster estimator and three estimators with individual-specific endogenous cut-off (the mean, median, minimum Hessian (FF)). Theoretical background and description of used estimators is presented in Appendix B.

The results reported in Table 8 for intrinsic motivation show only minor differences across models.[18] Column 1 reports coefficients from OLS estimator with

[17] The results hold also when we look on all classes specializing on any course (languages, sports, music and arts). Results available upon request.

[18] Estimates for other outcomes variables are not included but show similar patterns. Dummy variables for motivation of teachers, indicator variables for imputed values and dummy variables for each subjects are not reported in Table 8. However, results are consistent with the results presented in this section.

fixed-effects using four point scale. Coefficients of linear model can be interpreted as marginal effects. Even though ordered models cannot be compared in terms of magnitude (Column 2-5), coefficients have the same signs and resemble in majority of significance levels.[19] Moreover, relative magnitudes of coefficients across models are also roughly the same. This is especially important for coefficients in variables of our interest, teaching practices.

5 Conclusion

Taken together, we show that modern teaching practices have a significant impact on socio-emotional skills in technical subjects with no adverse effects on test scores. Standard practices make boys demotivated and lead to lower self-confidence. On the other hand, high-achieving girls benefit from standard practices. Importantly, positive effects of modern practices are higher than standard practices for both genders.

We believe that our findings can shed light on current public debate about teaching methods and their effectiveness for development of children. In spite of higher focus on socio-emotional skills among researchers and policy makers in recent years, we know very little how teaching methods used every day in classrooms affect them. This is quite surprising since changes in composition of lessons could be very cheap and scalable. Moreover, recent studies (Algan et al., 2013, Bietenbeck, 2014) suggest that modern practices such as working in small groups account for increase in skills not captured in test scores such as social capital or application of gained knowledge.

Our results advocate for wider inclusion of modern teaching practices in lessons. However, we have to emphasize that our results should be interpreted with caution. First, since we cannot rule out all confounding factors we refrain from interpreting our results as causal. Second, the results are based on the analysis of science

[19]Complicated procedures for executing FE ordered models prevent us from using exactly the same clustering procedure appropriate for TIMSS dataset. This could be the reason for a few dissimilarities in significance levels across models.

subjects and math and the relationship may differ for other subjects. Third, the positive impacts may vary for composition of lesson favouring modern practices. For example, it could be argued that marginal benefits of modern practices are high in CEE countries because they are used for a relatively short time and students may perceive them as more interesting and motivating. Therefore, more research is needed to confirm our findings for other countries and to specifically disentangle elements of modern practices which are essential for development of socio-emotional skills.

References

Algan, Y., Cahuc, P., and Shleifer, A. (2013). Teaching Practices and Social Capital. *American Economic Journal: Applied Economics*, 5(3):189–210.

Aslam, M. and Kingdon, G. (2011). What can teachers do to raise pupil achievement? *Economics of Education Review*, 30(3):559–574.

Baetschmann, G., Staub, K. E., and Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):685–703.

Beilock, S. L., Gunderson, E. a., Ramirez, G., and Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, 107(5):1860–3.

Bertrand, M. and Pan, J. (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics*, 5(1):32–64.

Bietenbeck, J. (2014). Teaching practices and cognitive skills. *Labour Economics*, 30:143–153.

Borghans, L., Duckworth, A. L., Heckman, J. J., and ter Weel, B. (2008). The economics and psychology of personality traits. *The Journal of Human Resources*, 43(4):972–1059.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies*, 47(1):1980.

Cornwell, C., Mustard, D. B., Parys, J. V., Cornwell, C., Mustard, D. B., and Parys, J. V. (2013). Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments : Evidence from Primary School Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments Evidence from Primary School. *Journal of Human Resources*, 48(1):236–264.

Cunha, F. and Heckman, J. J. (2006). Interpreting the Evidence on Life Cycle Skill Formation. *Handbook of the Economics of Education*, 1:697–812.

Dickerson, A., Hole, A. R., and Munford, L. A. (2014). The relationship between well-being and commuting revisited: Does the choice of methodology matter? *Regional Science and Urban Economics*, 49:321–329.

Dubey, P. and Geanakoplos, J. (2010). Grading exams: 100,99,98,... or A,B,C? *Games and Economic Behavior*, 69(1):72–94.

Ferrer-i Carbonell, A. and Frijters, P. (2004). How Important is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal*, 114(497):641–659.

Fiszbein, A. (2001). *Decentralizing Education in Transition Societies*. Number World Bank Publications.

Frey, B. S. and Stutzer, A. (2000). Happiness , Economy and Institutions. *The Economic Journal*, 110(466):918–938.

Gneezy, U., Meier, S., and Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4):191–210.

Golsteyn, B. H. H., Grönqvist, H., and Lindahl, L. (2014). Adolescent time preferences predict lifetime outcomes. *Economic Journal*, 124(580):F739–F761.

Gorard, S., Rees, G., and Salisbury, J. (1999). Reappraising the Apparent Underachievement of Boys at School. *Gender and Education*, 11(4):441–454.

Heckman, J., Moon, S. H., Pinto, R., Savelyev, P., and Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics*, 1(1):1–46.

Heckman, J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.

Houtte, M. V. (2004). Why boys achieve less at school than girls: the difference between boys' and girls' academic culture. *Educational Studies*, 30(2):159–173.

Hoxby, C. M. H. U. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. *National Bureau of Economic Research*, No. w7867.

Jacob, B. A. (2002). Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education review*, 21(6):589–598.

Joyce, T., Crockett, S., Jaeger, D. A., Altindag, O., and O'Connell, S. D. (2015). Does classroom time matter? *Economics of Education Review*, 46:64–77.

Kane, T. J., Taylor, E. S., Tyler, J. H., and Wooten, A. L. (2011). Identifying Effective Classroom Practices Using Student Achievement Data. *Journal of Human Resources*, 46(3):587–613.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., and Borghans, L. (2014). Fostering and Measuring Skills. *National Bureau of Economic Research,*, No. w19656.

Klaveren, C. V. (2011). Lecturing style teaching and student performance. *Economics of Education Review*, 30(4):729–739.

Koch, A., Nafziger, J., and Nielsen, H. S. (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization*, 115:3–17.

Lavy, V. (2015). What makes an effective teacher? Quasi-experimental evidence. *CESifo Economic Studies*.

Meece, J. L., Wigfield, A., and Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, 82(1):60–70.

Mischel, W., Shoda, Y., and Rodriguez, M. I. (1989). Delay of gratification in children. *Science (New York, N.Y.)*, 244(4907):933–8.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., and Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7):2693–8.

Mullis, I., Martin, M., Foy, P., and Arora, A. (2012a). TIMSS International Results in Science. Technical report.

Mullis, I. V. S., Martin, M. O., Foy, P., and Arora, A. (2012b). TIMSS 2011 International Results in Mathematics. Technical report, Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

Neyman, J. and Scott, E. L. (1948). Consistent Estimates Based on Partially Consistent Observations. *Econometrica*, 16(1):1–32.

Niederle, M. and Vesterlund, L. (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspectives*, 24(2):129–144.

OECD (2015). *The ABC of Gender Equality in Education*, volume PISA.

OECD (2016). PISA 2015 Results (Volume I) Excellence and Equity in Education. Technical report, OECD Publishing, Paris.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2):247–252.

Schwerdt, G. and Wuppermann, A. C. (2011). Is traditional teaching really all that bad? A within- student between-subject approach. *Economics of Education Review*, 30(2):365–379.

Spencer, S., Steele, C., and Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, 35(1):4–28.

Staiger, D. O. and Rockoff, J. E. (2010). Searching for Effective Teachers with Imperfect Information. *Journal of Economic Perspectives*, 24(3):97–118.

Tangney, J. P., Baumeister, R. F., Boone, A. L., Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success. *Journal of personality*, 72(2):271–324.

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

Table 1: Descriptive statistics
Teaching practices

	Mean	SD
Standard teaching practices		
Total average	0.64	(0.07)
Mathematics	0.67	(0.07)
Physics	0.65	(0.08)
Biology	0.61	(0.09)
Chemistry	0.68	(0.07)
Earth science	0.60	(0.08)
Difference (p-value)	0.00	
Modern teaching practices		
Total average	0.44	(0.06)
Mathematics	0.39	(0.08)
Physics	0.48	(0.08)
Biology	0.45	(0.08)
Chemistry	0.49	(0.08)
Earth science	0.39	(0.07)
Difference (p-value)	0.00	
Other activities	0.28	(0.08)

Notes: Class-aggregated indices leave out each student's own observation. Index expresses the share of lesson devoted to a particular teaching practice. Student weights are used. Chi-square multivariate test is used to compare means. Standard deviation in parentheses.

Table 2: Descriptive statistics
Teacher and class characteristics

Variable	Mathematics		Physics		Biology		Chemistry		Earth science		Difference
Teacher characteristics	207 teachers		206 teachers		208 teachers		203 teachers		202 teachers		p-value
Age:											
Under 25	0.01	(0.09)	0.00	(0.09)	0.00	(0.06)	0.02	(0.16)	0.01	(0.13)	0.30
25 to 29	0.12	(0.32)	0.09	(0.29)	0.12	(0.33)	0.07	(0.26)	0.21	(0.41)	0.00
30 to 39	0.21	(0.40)	0.21	(0.41)	0.12	(0.32)	0.24	(0.42)	0.19	(0.39)	0.00
40 to 49	0.33	(0.47)	0.32	(0.46)	0.23	(0.42)	0.23	(0.42)	0.23	(0.42)	0.02
50 to 59	0.26	(0.43)	0.20	(0.40)	0.38	(0.48)	0.31	(0.46)	0.23	(0.42)	0.00
60 or older	0.07	(0.26)	0.15	(0.35)	0.12	(0.33)	0.11	(0.31)	0.09	(0.29)	0.15
Female	0.77	(0.41)	0.56	(0.49)	0.83	(0.37)	0.85	(0.35)	0.61	(0.48)	0.00
Experience:											
0-2 years	0.03	(0.19)	0.06	(0.23)	0.07	(0.26)	0.08	(0.28)	0.12	(0.32)	0.02
3-5 years	0.10	(0.31)	0.09	(0.29)	0.06	(0.24)	0.06	(0.25)	0.11	(0.32)	0.25
6 or more years	0.85	(0.35)	0.84	(0.36)	0.86	(0.34)	0.84	(0.36)	0.76	(0.42)	0.11
University degree	0.96	(0.18)	0.93	(0.24)	0.96	(0.19)	0.96	(0.19)	0.92	(0.26)	0.24
Motivation (1-4 scale)											
Discuss concepts with others	1.81	(0.63)	1.89	(0.74)	1.77	(0.73)	1.83	(0.73)	1.75	(0.75)	0.39
Preparation of materials	1.99	(0.82)	2.10	(0.99)	1.95	(0.85)	1.96	(0.83)	1.95	(0.85)	0.45
Visits in other classes	1.13	(0.38)	1.15	(0.35)	1.12	(0.32)	1.11	(0.32)	1.14	(0.39)	0.80
Informal visits	1.08	(0.29)	1.16	(0.38)	1.12	(0.34)	1.13	(0.37)	1.10	(0.33)	0.15
Further development											
Subject content course	0.52	(0.50)	0.59	(0.49)	0.63	(0.48)	0.72	(0.44)	0.51	(0.50)	0.00
Pedagogy course	0.51	(0.50)	0.54	(0.49)	0.39	(0.49)	0.54	(0.49)	0.32	(0.46)	0.00
Curriculum improvement course	0.39	(0.49)	0.31	(0.46)	0.30	(0.45)	0.40	(0.49)	0.22	(0.41)	0.00
Subject related to IT	0.54	(0.49)	0.61	(0.48)	0.58	(0.49)	0.57	(0.49)	0.52	(0.50)	0.40
Critical thinking course	0.29	(0.45)	0.37	(0.48)	0.30	(0.45)	0.30	(0.46)	0.30	(0.46)	0.43
Student evaluation course	0.22	(0.41)	0.23	(0.42)	0.23	(0.42)	0.22	(0.41)	0.20	(0.40)	0.94
Class characteristics											
Class size	24.29	(4.20)	24.36	(4.03)	24.53	(4.13)	24.45	(4.12)	24.40	(4.00)	0.98
Minutes in class	203.12	(30.03)	116.60	(51.60)	91.18	(24.88)	100.41	(30.45)	92.98	(40.95)	0.00

Notes: Teacher probability weights are used to calculate all descriptive statistics. Variables on motivation range from 1-4 (1 - never or almost never, 2 - 2 or 3 times per month, 3 - 1-3 times a week, 4 - daily or almost daily), other variables from 0 to 1. Chi-square multivariate test is used to compare means. Standard deviation in parentheses.

Table 3: Teaching practices, socio-emotional skills and test scores

	Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
			University	Job	
	(1)	(2)	(3)	(4)	(5)
standard practices	0.108	0.202	0.400***	0.200	-0.393***
	(0.161)	(0.223)	(0.120)	(0.128)	(0.146)
modern practices	0.175	2.460***	0.396**	0.690***	1.140***
	(0.140)	(0.214)	(0.156)	(0.132)	(0.170)
Teacher female	-0.014	-0.034	-0.033	-0.046**	0.018
	(0.018)	(0.044)	(0.020)	(0.021)	(0.031)
Age					
25-29	0.017	0.008	-0.006	0.075	0.066
	(0.078)	-0.117	(0.054)	(0.049)	(0.137)
30-39	-0.023	-0.191	-0.048	0.035	-0.051
	(0.087)	(0.121)	(0.066)	(0.054)	(0.139)
40-49	-0.039	-0.186	-0.071	0.027	-0.059
	(0.093)	(0.129)	(0.069)	(0.055)	(0.142)
50-59	-0.033	-0.281**	-0.085	0.028	-0.108
	(0.094)	(0.121)	(0.065)	(0.059)	(0.136)
$>$ 60	-0.008	-0.231*	-0.058	0.026	-0.055
	(0.101)	(0.132)	(0.069)	(0.058)	(0.147)
Experience					
3-5 years	-0.02	0.151**	-0.049	-0.088**	0.035
	(0.030)	(0.077)	(0.042)	(0.041)	(0.070)
$>$ 5 years	0.027	0.227***	0.038	-0.029	0.051
	(0.047)	(0.064)	(0.051)	(0.041)	(0.058)
University diploma	0.043	-0.177***	-0.052	-0.027	-0.028
	(0.028)	(0.065)	(0.049)	(0.034)	(0.059)
Number of min./week	0.0003	0.0004	-0.0001	-0.0003	0.0004
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Class size	-0.0013	-0.0008	-0.0002	0.001	0.001
	(0.0014)	(0.003)	(0.002)	(0.001)	(0.002)
Further development					
Subject content course	-0.030*	-0.026	0.038**	0.008	-0.03
	(0.016)	(0.033)	(0.019)	(0.019)	(0.026)
Pedagogy course	0.017	0.011	-0.019	0.006	0.01
	(0.017)	(0.032)	(0.020)	(0.017)	(0.020)
Curriculum improvement course	-0.022	-0.04	0.002	0.012	-0.036
	(0.017)	(0.033)	(0.020)	(0.020)	(0.031)
Subject related to IT	-0.011	0.016	0.003	-0.016	-0.005
	(0.016)	(0.033)	(0.019)	(0.017)	(0.026)
Critical thinking course	-0.002	-0.013	-0.006	-0.004	-0.012
	(0.023)	(0.039)	(0.023)	(0.018)	(0.028)
Student evaluation course	0.014	0.064	0.004	-0.019	0.059*
	(0.028)	(0.040)	(0.019)	(0.020)	(0.030)
STP = MTP (p-value)	0.794	0	0.988	0.030	0
Observations	22,633	22,633	22,633	22,633	22,633
R-squared	0.808	0.428	0.521	0.502	0.435

Notes: All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Table reports p-values from hypothesis testing if standard teaching practices coefficient equals modern teaching practices coefficients.

Table 4: Teaching practices and gender effects

	Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
			University	Job	
	(1)	(2)	(3)	(4)	(5)
Panel A: Girls					
Standard practices	0.203	0.826***	0.526***	0.373**	-0.081
	(0.205)	(0.302)	(0.148)	(0.188)	(0.183)
Modern practices	0.155	2.196***	0.393**	0.507***	0.910***
	(0.132)	(0.270)	(0.188)	(0.171)	(0.197)
Observations	11,053	11,053	11,053	11,053	11,053
R-squared	0.819	0.424	0.527	0.514	0.442
Panel A1: Low-achieving					
Standard practices		0.432	0.291	0.0735	-0.472*
		(0.364)	(0.270)	(0.316)	(0.284)
Modern practices		2.253***	0.223	0.512**	0.816**
		(0.356)	(0.290)	(0.225)	(0.325)
Panel A2: High-achieving					
Standard practices		1.067***	0.799***	0.829***	0.261
		(0.365)	(0.242)	(0.266)	(0.259)
Modern practices		2.070***	0.637**	0.567**	0.938***
		(0.345)	(0.250)	(0.255)	(0.272)
Panel B: Boys					
Standard practices	0.013	-0.413*	0.257	0.007	-0.726***
	(0.160)	(0.217)	(0.178)	(0.154)	(0.180)
Modern practices	0.178	2.754***	0.394*	0.892***	1.376***
	(0.197)	(0.244)	(0.202)	(0.197)	(0.202)
Observations	11,580	11,580	11,580	11,580	11,580
R-squared	0.803	0.458	0.536	0.520	0.439

Notes: All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Central and Eastern European countries - teaching practices

	Standard practices	Modern practices	Other practices
Bulgaria	0.63	0.47	0.28
	(0.11)	(0.12)	(0.1)
Hungary	0.56	0.45	0.32
	(0.09)	(0.1)	(0.08)
Lithuania	0.63	0.38	0.32
	(0.1)	(0.09)	(0.08)
Romania	0.58	0.47	0.29
	(0.13)	(0.11)	(0.11)
Russia	0.7	0.53	0.37
	(0.12)	(0.09)	(0.07)
Slovenia	0.66	0.5	0.32
	(0.08)	(0.09)	(0.08)
Ukraine	0.73	0.55	0.38
	(0.08)	(0.08)	(0.06)

Notes: Class-aggregated indices leave out each student's own observation. Index expresses share of lesson devoted to a particular teaching practice. Student weights are used.

Table 6: Central and Eastern European countries

		Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
				University	Job	
		(1)	(2)	(3)	(4)	(5)
Pooled data	Standard practices	0.072	0.384***	0.232***	0.146*	-0.062
		(0.110)	(0.086)	(0.081)	(0.083)	(0.084)
	Modern practices	0.331***	1.279***	0.500***	0.607***	0.321***
		(0.128)	(0.107)	(0.094)	(0.091)	(0.103)
	R-squared	0.717	0.413	0.516	0.498	0.427
Bulgaria	Standard practices	0.092	1.097***	0.792***	1.074***	0.185
		(0.185)	(0.218)	(0.217)	(0.230)	(0.180)
	Modern practices	0.134	0.802***	0.321	0.178	-0.285
		(0.202)	(0.197)	(0.196)	(0.160)	(0.191)
	R-squared	0.807	0.508	0.575	0.609	0.520
Hungary	Standard practices	0.249**	-0.005	0.068	0.063	-0.264
		(0.109)	(0.182)	(0.144)	(0.151)	(0.168)
	Modern practices	0.103	1.847***	0.422**	0.506***	0.749***
		(0.081)	(0.189)	(0.165)	(0.124)	(0.145)
	R-squared	0.808	0.431	0.574	0.521	0.493
Lithuania	Standard practices	0.202*	0.719***	0.634***	0.407***	0.052
		(0.109)	(0.173)	(0.126)	(0.123)	(0.141)
	Modern practices	0.115	0.863***	0.526***	0.592***	0.071
		(0.096)	(0.191)	(0.117)	(0.119)	(0.159)
	R-squared	0.804	0.405	0.544	0.557	0.471
Romania	Standard practices	0.164	0.726***	0.364**	0.302**	0.096
		(0.157)	(0.178)	(0.158)	(0.138)	(0.146)
	Modern practices	0.143	1.056***	0.352**	0.419***	0.190
		(0.164)	(0.154)	(0.146)	(0.128)	(0.166)
	R-squared	0.796	0.479	0.559	0.581	0.455
Russia	Standard practices	-0.140	0.572***	0.381**	0.355**	-0.334**
		(0.119)	(0.199)	(0.156)	(0.138)	(0.133)
	Modern practices	0.291**	1.019***	0.355*	0.486***	0.498***
		(0.140)	(0.200)	(0.185)	(0.163)	(0.166)
	R-squared	0.809	0.451	0.584	0.554	0.480
Slovenia	Standard practices	0.285	0.617**	0.267	0.145	0.220
		(0.201)	(0.272)	(0.239)	(0.230)	(0.255)
	Modern practices	0.098	1.313***	0.581***	0.697***	0.347**
		(0.154)	(0.228)	(0.144)	(0.143)	(0.156)
	R-squared	0.804	0.508	0.565	0.527	0.520
Ukraine	Standard practices	0.05	1.158***	0.501**	0.005	0.138
		(0.203)	(0.242)	(0.219)	(0.221)	(0.235)
	Modern practices	0.443***	2.046***	0.860***	1.089***	0.499***
		(0.126)	(0.206)	(0.172)	(0.160)	(0.191)
	R-squared	0.795	0.496	0.572	0.576	0.502

Notes: Pooled data include all 7 CEE countries below. Number of observations: Pooled 109,182, Bulgaria 12,249, Hungary 19,127, Slovenia 14,903, Lithuania 18,622, Romania 18,823, Russia 20,969, Romania 18,823, Russia 20,969, Ukraine 19,392. All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Robustness checks

	Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
			University	Job	
	(1)	(2)	(3)	(4)	(5)
Alternative definitions of teaching practices					
Panel A1: Definition Alt. 1					
Standard practices	0.131	0.128	0.382***	0.117	-0.457***
Modern practices	0.218	2.639***	0.466***	0.740***	1.272***
Observations	22,392	22,392	22,392	22,392	22,392
R-squared	0.809	0.429	0.523	0.504	0.438
Panel A2: Definition Alt. 2					
Standard practices	-0.0177	0.231	0.346**	0.129	-0.451***
Modern practices	0.201	2.469***	0.444***	0.727***	1.125***
Observations	22624	22624	22624	22624	22624
R-squared	0.807	0.428	0.522	0.503	0.435
Course specific teaching practices					
Panel B1: Classes with a focus on technical subjects					
Standard practices	-0.546	-1.261	-0.859	-0.465	-0.664
Modern practices	0.784*	2.617**	0.862	0.307	1.454
Observations	1,803	1,803	1,803	1,803	1,803
R-squared	0.779	0.438	0.527	0.515	0.424
Restricted sample for special classes					
Standard practices	0.0732	0.224	0.386***	0.193	-0.438***
Modern practices	0.149	2.429***	0.365**	0.693***	1.141***
Observations	20,830	20,830	20,830	20,830	20,830
R-squared	0.808	0.429	0.522	0.502	0.437
Panel B2: Grammar schools					
Standard practices	-0.293	-1.504	-0.147	0.249	-2.124**
Modern practices	-0.147	0.819	0.343	0.686	-0.445
Observations	1,967	1,967	1,967	1,967	1,967
R-squared	0.713	0.388	0.420	0.427	0.419
Restricted sample for grammar schools					
Standard practices	0.0963	0.240	0.397***	0.148	-0.362**
Modern practices	0.150	2.512***	0.417**	0.719***	1.186***
Observations	20,666	20,666	20,666	20,666	20,666
R-squared	0.783	0.435	0.531	0.510	0.440

Notes: Panel A uses alternative definitions from Table A1. Panel B1 splits the sample classes with and without technical specialization (math, ICT, science). All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Comparison of linear and order models

Dependent var.: Intrinsic motivation	OLS	BUC	Mean	Median	FF
	(1)	(2)	(3)	(4)	(5)
Standard practices	0.202	0.385	0.202	0.0132	-0.135
	(0.223)	(0.297)	(0.278)	(0.277)	(0.365)
Modern practices	2.460***	5.651***	5.525***	5.387***	6.287***
	(0.214)	(0.299)	(0.279)	(0.277)	(0.382)
Teacher female	-0.034	-0.084*	-0.129***	-0.107**	-0.129**
	(0.044)	(0.047)	(0.043)	(0.044)	(0.058)
Age					
25-29	0.008	0.075	0.213	0.224	0.398*
	-0.117	(0.169)	(0.154)	(0.155)	(0.217)
30-39	-0.191	-0.426**	-0.220	-0.186	-0.028
	(0.121)	(0.184)	(0.167)	(0.168)	(0.233)
40-49	-0.186	-0.387**	-0.190	-0.140	0.002
	(0.129)	(0.187)	(0.169)	(0.170)	(0.237)
50-59	-0.281**	-0.621***	-0.365**	-0.325*	-0.221
	(0.121)	(0.187)	(0.170)	(0.171)	(0.237)
> 60	-0.231*	-0.480**	-0.263	-0.242	-0.182
	(0.132)	(0.199)	(0.179)	(0.179)	(0.247)
Experience					
3-5 years	0.151**	0.370***	0.218**	0.211**	0.195
	(0.077)	(0.0980)	(0.0897)	(0.0904)	(0.122)
> 5 years	0.227***	0.552***	0.428***	0.364***	0.472***
	(0.064)	(0.107)	(0.099)	(0.010)	(0.138)
University diploma	-0.177***	-0.427***	-0.347***	-0.334***	-0.546***
	(0.065)	(0.096)	(0.091)	(0.092)	(0.123)
Number of min./week	0.0004	0.0008*	0.0008*	0.0011***	0.0026***
	(0.0003)	(0.0005)	(0.0004)	(0.0004)	(0.0006)
Class size	-0.0008	-0.0013	0.0007	0.0008	-0.0055
	(0.003)	(0.0040)	(0.0036)	(0.0036)	(0.0049)
Further development					
Subject content course	-0.026	-0.041	-0.022	-0.061	-0.001
	(0.033)	(0.045)	(0.043)	(0.043)	(0.059)
Pedagogy course	0.011	0.018	-0.008	0.011	0.066
	(0.032)	(0.046)	(0.043)	(0.042)	(0.057)
Curriculum improvement course	-0.04	-0.093**	-0.095**	-0.075*	-0.111*
	(0.033)	(0.046)	(0.044)	(0.043)	(0.059)
Subject related to IT	0.016	0.044	0.072*	0.055	0.004
	(0.033)	(0.041)	(0.038)	(0.039)	(0.053)
Critical thinking course	-0.013	-0.038	-0.062	-0.041	-0.049
	(0.039)	(0.049)	(0.046)	(0.045)	(0.063)
Student evaluation course	0.064	0.139***	0.121**	0.122**	0.067
	(0.040)	(0.052)	(0.048)	(0.048)	(0.066)
Observations	22,633	40,073	21,092	21,087	12089
R-squared	0.428				

Notes: Dependent variable is intrinsic motivation in all columns. Dependent variable is on 4 point scale (Agree a lot, agree a little, disagree a little, disagree a lot). Apart from controls in the table it is controlled for motivation of teachers, subjects dummy variables and imputed indicators. Column 1 estimated with OLS, other with ordered-logit fixed-effects models (Blow-up and cluster estimator (column 2) and 3 estimators with endogenous cut-off (Mean, median and Hessian)). Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A1: Definitions of teaching practices

Original			
Standard practices	1		We listen to the teacher give a lecture-style presentation
	2		We memorize formulas and procedures
	3		We work problems on our own
Modern practices	1		We explain our answers
	2		We relate what we are learning in *subject* to our daily lives
	3	Math	We work together in small groups
		Science	We work in small groups on an experiment or investigation
Alternative 1			
The fourth practice is added			
Standard practices	4	Math	We write equations and functions to represent relationships
		Science	We use scientific formulas and laws to solve problems
Modern practices	4	Math	We decide on our own procedures for solving complex problems -
		Science	We design or plan an experiment or investigation
Alternative 2			
Standard teaching practice "Work alone" replaced			
Standard practices	3	Math	We write equations and functions to represent relationships
		Science	We read our science textbooks and other resource materials

Appendix

Appendix A - Additional results

Table A2: Interaction of test scores and socio-emotional skills

	Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
			University	Job	
	(1)	(2)	(3)	(4)	(5)
Standard practices	0.096	0.184	0.391***	0.189	-0.407***
	(0.161)	(0.222)	(0.119)	(0.128)	(0.146)
Modern practices	0.030	2.428***	0.383**	0.672***	1.114***
	(0.137)	(0.214)	(0.156)	(0.132)	(0.171)
Test score		0.179***	0.077**	0.099***	0.145***
		(0.026)	(0.025)	(0.028)	(0.025)
Intrinsic motivation	0.059***				
	(0.008)				
Extrinsic - university	0.017				
	(0.010)				
Extrinsic - job	0.024**				
	(0.010)				
Self-confidence	0.031***				
	(0.008)				
Observations	22,633	22,633	22,633	22,633	22,633
R-squared	0.806	0.434	0.523	0.504	0.439

Notes: Column 1 reports coefficients from five regressions, test scores and one of socio-emotional outcome as a control. Coefficient for test scores are from regression with intrinsic motivation. Test scores coefficients in other regressions are also insignificant and range from 0.103 to 0.127 for SP and from 0.117 to 0.162 for MP. Similarly, R^2 ranges from 0.807 to 0.809. All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors. in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Interaction of subjects and socio-emotional skills

	Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
			University	Job	
	(1)	(2)	(3)	(4)	(5)
Standard practices	0.160	-0.376	0.063	0.062	-0.474**
	(0.212)	(0.261)	(0.194)	(0.204)	(0.213)
Modern practices	0.056	3.016***	0.749***	0.928***	1.112***
	(0.197)	(0.354)	(0.226)	(0.181)	(0.315)
Math x standard	-0.031	1.152***	0.672**	0.299	-0.367
	(0.354)	(0.383)	(0.335)	(0.351)	(0.347)
Physics x standard	-0.023	0.686*	0.502*	0.271	-0.355
	(0.289)	(0.382)	(0.264)	(0.233)	(0.309)
Biology x standard	0.078	0.607	0.110	-0.148	0.08
	(0.290)	(0.413)	(0.266)	(0.265)	(0.349)
Chemistry x standard	-0.310	0.463	0.565**	0.379	0.207
	(0.280)	(0.453)	(0.285)	(0.241)	(0.381)
Math x modern	0.036	-1.423***	-0.655**	-0.289	0.477
	(0.318)	(0.530)	(0.325)	(0.299)	(0.395)
Physics x modern	-0.159	-0.581	-0.520*	-0.305	0.059
	(0.329)	(0.591)	(0.280)	(0.276)	(0.457)
Biology x modern	0.309	-0.208	-0.204	-0.228	-0.199
	(0.330)	(0.543)	(0.349)	(0.354)	(0.436)
Chemistry x modern	0.487*	-0.496	-0.375	-0.393	-0.408
	(0.291)	(0.516)	(0.338)	(0.304)	(0.418)
Observations	22,633	22,633	22,633	22,633	22,633
R-squared	0.808	0.429	0.522	0.502	0.436
Controls	YES	YES	YES	YES	YES

Notes: Earth science is the omitted subject. All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors. in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Teaching practices separately and functional forms

	Test score	Intrinsic motivation	Extrinsic motivation		Self-confidence
			University	Job	
	(1)	(2)	(3)	(4)	(5)
Panel A					
Standard practices	0.170	1.096***	0.560***	0.456***	-0.0336
	(0.147)	(0.243)	(0.106)	(0.114)	(0.146)
Observations	23,570	23,327	23,464	23,489	23,323
R-squared	0.804	0.405	0.515	0.493	0.425
Panel B					
Modern practices	0.192	2.509***	0.545***	0.748***	-0.983***
	(0.126)	(0.190)	(0.129)	(0.107)	(0.161)
Observations	23,484	23,245	23,387	23,407	23,245
R-squared	0.805	0.423	0.515	0.496	0.429
Panel C					
Standard practices	0.646	-0.701	-0.489	-0.051	-1.357*
	(0.749)	(1.098)	(0.763)	(1.042)	(0.802)
Modern practices	-0.270	2.923***	0.021	0.741	0.225
	(0.527)	(0.989)	(0.533)	(0.601)	(0.812)
Standard practices^2	-0.423	0.719	0.732	0.202	0.810
	(0.606)	(0.908)	(0.606)	(0.772)	(0.654)
Modern practices^2	0.490	-0.505	0.434	-0.053	1.041
	(0.592)	(1.166)	(0.597)	(0.669)	(0.921)
Observations	22,633	22,633	22,633	22,633	22,633
R-squared	0.808	0.428	0.521	0.502	0.435

Notes: Panel A estimated for standard practices only and Panel B for modern practices only. Panel C added to the original specification practices squared. All regressions control for teacher and class characteristics from Table 2. Furthermore, regressions control for subject dummies and imputation indicators. Test score coefficients are estimated from five plausible values. Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B - Fixed-effects ordered logit models

This section defines fixed-effects ordered logit model and then reviews estimators that have been suggested in the literature based on (Baetschmann et al., 2015, Dickerson et al., 2014). The model considers latent variable y^*_{it} for individual i in time t (in our case in 5 subjects) to a vector of observable characteristics x_{it} and unobserved characteristics α_i, ϵ_{it}:

$$y^*_{it} = \beta x'_{it} + \alpha_i + \epsilon_{it}, \qquad i = 1, \ldots, N \quad t = 1, \ldots, T. \tag{B1}$$

α_i is invariant unobserved component which may or may not be correlated with x_{it}. Observed ordered variable y_{it} is related to the latent variable y^*_{it} in the following way:

$$y_{it} = k \quad \text{if} \quad \upsilon_{ik} < y^*_{it} \leq \upsilon_{ik+1}, \quad k = 1, \ldots, K \tag{B2}$$

and the thresholds υ_k are increasing ($\upsilon_{ik} \leq \upsilon_{ik+1} \quad \forall k$) with $\upsilon_{i1} = -\infty$ and $\upsilon_{iK+1} = \infty$. The fixed-effects ordered logit model assumes that ϵ_{it} are IID with logistic cumulative distribution function $\Lambda(\cdot)$ and the probability of observing outcome k for individual i in time (subject) t is

$$Pr(y_{it} = k | x_{it}, \alpha_i) = \Lambda(\upsilon_{ik+1} - \beta x'_{it} - \alpha_i) - \Lambda(\upsilon_{ik} - \beta x'_{it} - \alpha_i) \tag{B3}$$

Baetschmann et al. (2015) discuss two problems with estimation of maximum likelihood from the equation B3. The first is that only $\upsilon_{ik} - \alpha_i = \alpha_{ik}$ can be identified. The second is that under T asymptotics, estimation of α_{ik} is not consistent due to incidental parameters problem (Neyman and Scott, 1948). The bias of $\hat{\beta}$ can be substantial, especially in short panels.

The Blow-up and Cluster (BUC) estimator

Baetschmann et al. (2015) proposed an estimator which combines information from different cut-offs into a single likelihood function, yielding a one-step estimator of β. The model uses all $K-1$ cutoffs simultaneously and imposes restriction that $\beta^2 = \ldots = \beta^K$. It is implemented in the following way - every observation is replaced in the sample by $K-1$ copies of itself and each of the $K-1$ copies of the individual is dichotomized at a different cut-off point. The expanded sample can be then estimated using Chamberlain (1980) approach. Standard errors have to be clustered at the individual level since some observations are used multiple times. Baetschmann et al. (2015) show that BUC estimator is consistent and performs well even in small panels.

The Ferrer-i-Carbonell and Frijters (FF) estimator

Ferrer-i Carbonell and Frijters (2004) suggested an estimator that identifies a single and "optimal" cut-off point for each individual. The optimal cut-off is found by minimizing the Hessian matrix. It is done in practice at a preliminary estimate of $\hat{\beta}$. We also include in our analysis computationally simpler approach which chooses the cut-off at individual mean or median of y_{it}. However, Baetschmann et al. (2015) show that FF estimators are inconsistent since the procedure chooses the cut-off point endogenously.

IES Working Paper Series

2017

1. Petra Lunackova, Jan Prusa, Karel Janda: *The Merit Order Effect of Czech Photovoltaic Plants*
2. Tomas Havranek, Zuzana Irsova, Tomas Vlach: *Measuring the Income Elasticity of Water Demand: The Importance of Publication and Endogeneity Biases*
3. Diana Zigraiova, Petr Jakubik: *Updating the Long Term Rate in Time: A Possible Approach*
4. Vaclav Korbel, Michal Paulus: *Do Teaching Practices Impact Socio-emotional Skills?*

All papers can be downloaded at: http://ies.fsv.cuni.cz

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz http://ies.fsv.cuni.cz