Is private supplementary tutoring effective? A longitudinally detailed analysis of private tutoring quality in China

Yuhuan Zhang[1] · Yu Dang[1] · Yahan He[2] · Xiao Ma[3] · Lidong Wang[4,5]

Abstract

In general, private supplementary tutoring is considered an effective means of improving academic achievement by parents. However, previous studies have produced partial or inconclusive results regarding its effectiveness. Thus, the present study conducted a comprehensive analysis based on a specially designed longitudinal survey of private supplementary mathematics tutoring among middle school students in China. The hierarchical linear analysis did not show a considerably significant positive effect of regular private tutoring, even when such tutoring continued throughout the summer vacation and school semester. The data also allowed a deeper inquiry into the role of the quality, quantity, and applicability of private tutoring. In addition, the heterogeneous effect of private supplementary tutoring across students with different mathematics achievement levels was explored, with possible explanations including the different learning styles of students from different groups, the ceiling effect of the testing tool, and the different learning needs of different groups of students. These results suggest that parents should carefully select private tutoring for their children, and that the government should provide more comprehensive professional guidelines to regulate the industry.

Keywords Private supplementary tutoring · Mathematics achievement · Heterogeneous effect · Instructional quality · China

Introduction

Private supplementary tutoring (hereafter referred to as "PT") is widely known as "shadow education", since much of it mimics the regular school curriculum. It has also become more popular in recent decades (Bray 2009; Cole 2016; Liu 2012). For example, PT is very common in East Asian countries, especially those with highly competitive school systems and high-stakes examinations, and has now spread worldwide (Bray and Lykins 2012; Choi and Park 2016; Park et al. 2016; Wang and Guo 2017; Ömeroğulları et al. 2020). It has become a standard option for parents and students who want to advance ahead of their schools' curricula or compensate for a deficiency in an academic subject (Smyth 2009; Wang et al. 2018, 2019). According to the Chinese National Assessment for Education Quality of Mathematics (a government-authorized, standardized, national education-oriented survey), approximately 44% of grade 4 and 23% of grade 8 students participated in mathematics PT in 2015, respectively, with roughly 45% of them receiving weekly PT outside of school for more than 2 h per week (National Center of Education Quality 2018).

Previous international academic achievement assessments have suggested a clear advantage in favor of Chinese and other Eastern Asian students over their Western peers (Leung 2001; Mullis et al. 2004; Wang and Lin 2005). This advantage is potentially attributable to their efforts in both in-school learning and PT outside of school (Wang and Guo 2017; Wang et al. 2018, 2019; Bray et al. 2020). However, with PT consuming significant family resources, obvious questions for stakeholders include whether PT effectively increases students'

✉ Lidong Wang
  wanglidong@bnu.edu.cn

1   School of Mathematics and Statistics, Henan University, Kaifeng, China

2   School of Mathematical Sciences, Beijing Normal University, Beijing, China

3   Institute of Sport Sciences, Shanghai University of Sport, Shanghai, China

4   Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University, Beijing, China

5   International Center for Research in Mathematics Education, Beijing Normal University, Beijing, China

academic achievement (Bray 2014), especially in countries with high rates of academic achievement and PT enrollment such as China.

In regard to possible answers, various efforts have been made by scholars all over the world in a wide array of fields, including educational sociology, pedagogy, educational policy, and psychology (Liu 2012; Zeng and Zhou 2012; Zhang 2011; Lee 2013; Choi and Park 2016; Zhang and Liu 2016; Guill et al. 2019; Ömeroğulları et al. 2020; Yung 2020). However, perhaps because of the limitations of datasets, the majority of these studies have merely focused on whether PT is effective by comparing the achievements of tutored and non-tutored students (overall effect), and controlling for relevant covariates with different statistical methodologies. This has also been insufficient for providing abundant evidence or implications for both theory and practice, especially since the results have been mixed or inconclusive (Bray 2014; Choi and Park 2016; Wang and Guo 2017; Ömeroğulları et al. 2020).

Meanwhile, the effects of different types of PT are more complicated (Choi and Park 2016; Zhang and Liu 2016; Ömeroğulları et al. 2020), especially with the difficulty of access to data and lack of comprehensive data (Zhang 2013). Recently, several studies have used all available data to attempt more detailed analyses of which characteristics of PT instruction contribute to the enhancement of academic achievement (Chappell, Nunnery, Pribesh, and Hager 2011; Liu 2012; Lee 2013; Zhang and Liu 2016; Guill et al. 2019; Ömeroğulları et al. 2020; Yung 2020). Yet, the problem persists because the measures of PT instruction are still undeveloped. Also, the alignment of assessment and PT instruction could be a problem since the aforementioned studies merely used available achievement scores, they might not align with the contents or cognitive demand of regular school or PT instruction.

In this context, the present study conducted a comprehensive analysis on the effectiveness of PT participation on mathematics achievement based on a specially designed retrospective longitudinal survey project that continued throughout the summer vacation and the school semester, focusing on mathematics PT of middle school students in China. Mathematics was selected for three reasons: (1) it is the subject with the largest PT enrollment in China (Wang and Guo 2017); (2) it is a subject with specific characteristics (Wang et al. 2018, 2019); and (3) it is a difficult subject from students' perspectives.

Literature review

Effectiveness of PT

Measures of PT

Several measures have been applied in previous studies to determine the characteristics and effectiveness of PT. These efforts are ongoing. Some of the most significant studies have described the overall effect of PT using whether students received PT, as a dummy variable (Zhang 2013), and comparing the achievements of tutored and non-tutored students (Baker et al. 2001; Kuan 2011; Zhang 2013; Bray 2014; Byun 2014; Wang and Guo 2017; Guill et al. 2019; Ömeroğulları et al. 2020), after controlling for the covariates with different statistical methods. Although PT is thought (or expected) to be effective, existing studies have shown weak positive, non-significant (and sometimes negative) effects on mathematics achievement, especially those with longitudinal research designs that controlled for selection bias through statistical methods (Kuan 2011; Byun 2014; Zhang 2013; Guill et al. 2019; Ömeroğulları et al. 2020).

The question by Bray (2014)—"Does private supplementary tutoring work"? —has been criticized as too broad a research agenda to be meaningful to policy or practice. Recently, studies have increasingly focused on the comprehensive measures of PT, since the nature of PT embraces a much broader range of modes than regular schooling (Bray 2014), the forms of PT are quite diverse, and the quality of PT instruction tends to vary across different forms of PT and different groups of students (Zhang 2013; Ömeroğulları et al. 2020). PT can also be remedial or aim for enrichment (Wang and Guo, 2017), while PT sessions can be conducted during or outside the school semester (Bray et al. 2020; Zhang, Ma, and Wang 2020). Meanwhile, the instructional coherence (Chen and Li 2010) between PT and regular school instruction can be a problem, as can the correlation of tutoring contents with school examinations (Zhang, et al. 2020).

PT can also have short- or long-term effects (Lee 2013). For example, the preview of in-class knowledge during summer vacation might impact the learning in the following semester, which in turn, can affect student achievement at the end of the semester (Zhang et al. 2020). However, a detailed survey needs to be conducted on the different types of PT, instead of merely asking if students participated in PT over the past 3 years (Bray 2014). For instance, the qualifications of tutors and the forms of instructional organization can be diverse, ranging from university and regular school teachers to university students (Wang and Guo 2017; Ömeroğulları et al. 2020), which might impact the overall effectiveness of PT.

Additionally, since collecting comprehensive data to measure the quality of PT can be particularly challenging (Zhang 2011; Wang et al. 2018, 2019), more research is necessary, especially with a specialized (longitudinal) database (Zhang 2013), rather than partial (cross-sectional) data from large survey programs (Hu et al. 2015). Meanwhile, previous studies have generally focused on the amount of PT (Liu 2012), with little discussion on the (instructional) quality of PT (Liu 2012; Zhang 2013; Byun 2014; Guill et al. 2019).

Some recent studies have discussed the effect of PT in more detail. For example, Lee (2013) found that PT in secondary schools affected students' academic achievement in both the short and long term using the methods of ordinary least squares, instrumental variables, and propensity score matching. The results suggested that PT in middle school (on average) had positive short-term effects on students' academic achievement, but minimal long-term effects on their university entrance examination scores. In related studies, Guill et al. (2019) revealed no effects of PT instructional quality on students' scores in mathematics, based on the three dimensions of structure, challenge, and support in PT, while Ömeroğulları et al. (2020) did not find an overall positive effect of more qualified tutors on mathematics when controlling for prior knowledge, and motivational and socio-demographic variables. Moreover, Zhang and Liu (2016) used class size as a measure and found a positive statistical effect for a larger class size. Overall, the aforementioned studies discussed the effects of PT (in detail) and offered new insights, which inspired us to conduct further research. However, since the results of existing studies regarding the characteristics of PT are still limited and inconclusive, the present study provides a more comprehensive picture.

The statistical approach

In regard to a quantitative evaluation of the effectiveness of PT, various statistical approaches were applied in previous studies, the majority of which followed the research agenda of educational effectiveness. Typically, regression models were applied by controlling the selection bias of PT and the endogeneity, based on the notion that PT participation is determined by some exogenous variables that also decide the dependent variable (Zhang 2013), including ordinary least squares (OLS) regression analysis (Ömeroğulları et al. 2020) and the hierarchical linear model (HLM) (Zhang 2013). More complicated methods have also been applied such as propensity score matching (Kuan 2011; Zhang and Liu 2016). Additionally, Zhang (2013) developed two instrument variables (IV) to address the problem of endogeneity, and found that PT participation negatively impacts mathematics achievement without significance, while the OLS results indicated a positive effect without significance. In a related study, Zhao (2015) used a heteroskedasticity-based identification and estimation method, and found that PT expenditure includes a small but statistically significant effect on the mathematics scores of primary school students.

In the absence of experimental work, with the exception of Meyer and Van Klaveren (2013), which examined the effectiveness of extended day programs through a randomized field experiment, stronger data analysis on causal factors should be conducted on longitudinal research, rather than solely relying on correlation analysis (Bray 2014). Cross-sectional data have also been insufficient in identifying the causal impact of PT on students' academic achievement, while longitudinal studies have focused on the contribution of PT to students' achievement gains, instead of focusing on their achievement at a given time, though control for the pre-achievement might not solve the entire problem (Todd and Wolpin 2003).

In sum, the identification of the causal effect of PT remains a challenge, especially since previous studies have yet to establish a solid theoretical framework (Guill et al. 2019). Thus, the results of the existing literature should be regarded as somewhat incomplete.

Theoretical model regarding the effectiveness of PT

Although previous studies have tested the effectiveness of several aspects of PT in improving students' academic achievement, limited theoretical consideration has been given to a comprehensive description of its effectiveness, with some exceptions including the studies based on Carroll's (1963) model of school learning (Guill et al. 2019; Ömeroğulları et al. 2020). This particular model posits that learning is a function of the ratio of the time spent on learning to the time required to learn, with the latter depending on cognitive factors and instruction quality. In this regard, PT adds external time to students' learning and potentially enhances it, but cognitive factors and instruction quality decide the effectiveness of the external time as well as the time required. However, this model does not provide a theoretical framework for describing what type of PT instruction is effective for enhancing students' learning.

Dunkin and Biddle (1974) provided a well-known scheme for classifying types of variables in the research on teaching through the "process–product" paradigm, which comprises four dimensions: (1) presage variables (properties of the teachers that can be assumed, but also have an influence on the interactive phase of teaching, e.g., teachers' experience and in-service training); (2) context variables (variables that have direct effects on the instructional outcomes and/or conditions that determine the effects of the process variables on the product variables, e.g., the characteristics of the students, schools, communities, and classes); (3) process variables (properties of the interactive phase of instruction as well as the phase of instruction during which students and teachers interact with the academic content, e.g., the observed actions of teachers and students as well as the instructional content or materials); and (4) product variables (possible outcomes of teaching, e.g., cognitive and non-cognitive changes in the students). Other than the aforementioned process variables, the presage variables, such as instructional quality, the presage variables of PT, such as the qualification of tutors (Ömeroğulları et al. 2020), were applied in previous studies. Overall, this model incorporating teacher effect research

Table 1 Theoretical framework

Dimension	Sub-dimension
Instructional quality	
Presage variables	Teacher qualifications; Time point; Forms of instructional organization
Process variables	Instructional contents
Product variables	Mathematics achievement
Cognitive factors	
Context variables	Pre-mathematics achievement; Self-expectation of education; Family socio-economic status; Demographic characteristics
Time spent	
Intensity	Hours spent on PT; Average hours spent on PT per week

(Wang and Cao 2014; Wang et al. 2018, 2019) can potentially provide a framework for comprehensively analyzing the aspects of effective PT, since some of the variables have only been partially discussed in previous studies.

The present study

The purpose of the present study was to conduct a more comprehensive analysis of the characteristics and effectiveness of PT (under a systematic theoretical framework) to provide more usable evidence for educational theory and practice under a comprehensive theoretical framework. In this case, the characteristics included the time spent on PT (Liu 2012), the instructional quality of PT (Guill et al. 2019; Ömeroğulları et al. 2020), and the heterogeneous effect of PT (Campbell et al. 2003), with a focus on both presage and process variables (Dunkin and Biddle 1974) such as summer/semester PT, strength, quality of tutoring, instructional contents, instructional organization (Zhang and Liu 2016; Ömeroğulları et al. 2020; Zheng et al. 2020). The present study also focused on the short-term effect within a school semester (and the preceding summer vacation), while leaving the evaluation of long-term effects for future research. Moreover, the dataset in the present study were developed according to the theoretical framework, which included evaluating both regular school and PT instruction, with the latter as a supplement to the former.

Theoretical framework

The present study used Carroll's (1963) model of school learning as well as Dunkin and Biddle's (1974) paradigm to determine the aspects of the instructional quality of PT that impact students' mathematics achievement, as illustrated in Table 1. Several presage or process variables discussed in previous studies as well as new developed variables were included in the framework to explain the effect of PT on the product variable (mathematics achievement in school). The

effectiveness was also explained by the time spent on PT, according to Carroll's (1963) model.

As discussed earlier, some of the variables in previous studies were included in the model, such as the qualifications of tutors (Ömeroğulları et al. 2020), forms of instructional organization (Byun 2014; Zhang and Liu 2016), and intensity (Liu 2012; Ömeroğulları et al. 2020), with some adaptations. It should be noted that this framework was not a saturated one, since additional variables were developed to provide more comprehensive information for further research.

PT, which is different from regular school instruction (which follows the textbook), can potentially help students improve their learning by previewing/reviewing the learning contents in regular school instruction (Wang et al. 2018, 2019). However, determining which forms can benefit the students' mathematics learning require further evidence. Thus, the variable of instructional contents was designed to address this issue.

Perhaps due to the limitations of data collection, previous studies have typically investigated PT activities over the past year (or more), without detailed differentiation. However, since the alignment between the contents of PT and testing have not been thoroughly discussed (Zhang et al. 2020), it is important to match the PT with the achievement data and distinguish the long-term effects from the short-term ones (Lee 2013). Meanwhile, the different roles of PT during the summer vacation and school semester have not been discussed (Bray et al. 2020). Hence, the variable of time point was designed to address this issue. More specifically, the variable of PT effect could "work collaboratively" (Bray 2014) by coding the different types of tutoring (e.g., one-on-one, small class, online, etc.) as simply yes/no, rather than examining/controlling for either different durations or different qualities of tutoring.

Summer vacation PT may have an impact on PT participation in the following semester (Zhang et al. 2020). Hence, there should be a focus on different types of interaction variables simultaneously to solve the problem of confounding effect.

Typically, research on the teacher effect has explored the heterogeneous effect of teachers on diverse groups of students to obtain detailed information on the effectiveness of PT (Campbell et al. 2003; Choi and Park 2016; Zhang 2013). As for PT participation, achievement has been found to be one of the most important factors (Byun et al. 2018; Zhang et al. 2020), and it is necessary to test). Therefore, the hypothesis that students from different achievement groups benefit from PT differently was tested.

Research questions

The present study focused on the following research questions

1. Does participating in PT, including certain combinations of PT during the summer vacation and school semester, have an overall effect on students' mathematics achievement?
2. Is the effectiveness of PT heterogeneous across students with different mathematics achievement?
3. According to the theoretical framework, do different forms of PT, decided by presage variables, process variables, and intensity of PT, explain the effectiveness of PT in increasing students' mathematics achievement?

Data

Participants and procedure

The dataset in the present study was developed to comprehensively evaluate the effectiveness of PT. Data collection occurred in Kaifeng, a typical medium-sized Chinese city located in central China, and placed at the mid-level of the country's economic and educational development. With the help of the Kaifeng Education Bureau, the data were collected through two processes: a questionnaire survey and administrative data collection, with the latter including the school-reported final examination scores for the semester.

For the questionnaire surveys, the targeted participants were grade 8 students, who completed the survey 1 week before the final examination in December 2019. Given that different schools had different quality levels, stratified targeted sampling was applied. All of the public schools were also divided into three levels under the supervision of the head of the department of mathematics and, after which five representative schools of different quality levels were recommended. Overall, a total of 2645 students were included.

Moreover, the students completed the questionnaires independently in class, under the supervision of the teacher in charge. The final examination was unified across the entire city by the Head of the Department of Mathematics. Due to the missing achievement data, three school levels were included in the final analysis (high, middle, and low). The final sample size was 2274, with the students' proportions of the three school levels at 37.3%, 42.7%, and 20%, respectively. However, after removing the samples with logically contradictory answers, a total of 1988 remained.

Finally, the EM imputation method in SPSS 22.0 was used to impute the remaining missing data, after which 1988 was the complete sample data (Leech et al. 2014). In this case, EM is a numerical algorithm that can be used to maximize likelihood in a wide variety of missing-data models (Dempster et al. 1977). It is also an iterative optimization strategy that is divided into two steps. In the expectation step, the expected log-likelihood is taken over the variables with missing data, under the given observation data and the current parameter estimation. In the maximization step, the expected log-likelihood is maximized to update the parameter value. The two steps are then alternately performed until convergence.

Measures

Dependent variables

The dependent variables were the school mathematics scores of the grade 8 students. The testing was designed by the Head of the Department of Mathematics to assess students' mathematics learning in the first semester of grade 8, thereby avoiding errors caused by inconsistent measuring tools (Zhang et al. 2020). The total score was 100, which was standardized into the z-score.

PT participation

The questionnaires mainly asked the students retrospectively (at two different time points) about their mathematics PT participation, summer vacation, and school semester. The students were also asked about their PT intensity (e.g., never; occasionally; 1 h per week on average; 2 h per week on average; 3 h per week on average; and 4 h or more per week on average). This information was then integrated into a PT variable so that the status of the PT could be divided into no PT, occasional PT, and regular PT. It should be noted that, strictly speaking, "regular PT" means receiving PT with significant intensity, rather than receiving PT every week. However, according to experience, PT during the semester is typically given (regularly) every one or two weeks, while PT during the summer vacation is typically given (regularly or intensively) over the course of several days.

The status of PT participation at the two time points was also transformed into eight interaction dummy variables (with no PT in both the summer vacation and school semester as the reference group), constituting the saturation model.

Table 2 Definitions and measures of the variables

Variable	Definitions and measures
Dependent variable	
Mathematics achievement	Mathematics achievement scores at the end of grade 8 (z-score, total score: 100)
Independent variables	
Participation	No PT, occasional PT, regular PT
Instruction contents	Preview, review, other
Teacher qualifications	University student, regular school teacher, cram school, university teacher, etc.
Time point	School semester, summer vacation
Average hours spent on PT per week	Never, occasionally, average 1 h per week, average 2 h per week, average 3 h per week, and average 4 h or more per week[a]
Hours spent on PT	18* average hours spent in PT per week in the semester + 6* average hours spent in PT per week in the vacation
Forms of instructional organization	One-to-one tutoring, class tutoring, and online tutoring
Covariates	
Prior knowledge	Mathematics achievement scores at the end of grade 7 (total score: 100)
Gender	Male = 0; Female = 1
Family economy	Five categories: very poor (reference group), relatively poor, medium, relatively rich, and very rich
Mother's education	Five categories: primary school or less; junior high school (reference group); technical secondary school, vocational high school, senior high school; junior college, undergraduate; masters; PhD
Educational expectations	Four categories: senior high school or lower (reference group); junior college; undergraduate; masters; PhD

[a]"Occasional" and "never" were coded as 0, while 1, 2, 3, and 4 h (or more) per week were coded as 1, 2, 3, and 4, respectively

In addition, the participants were asked to provide detailed information about the contents of the tutorials (preview, review, other), the form of instructional organization (one-to-one tutoring, class tutoring or online tutoring), and the qualification of the tutors (undergraduate or post-graduate university students, regular school teachers, cram school teachers, university teachers, etc.) at the two time points. It should be noted that in the few studies involving tutors, there was not a well-accepted (official) standard for the qualification of the tutors, and the aforementioned characteristics were typically used in existing studies (e.g., Ömeroğulları et al. 2020), as an initial exploration of such qualifications based on their career backgrounds.

Regarding PT intensity, the students had 6 weeks for the summer vacation and 18 weeks for the semester. Thus, the weight of the average hours spent in PT per week at the two time points was set and integrated into the variable of hours spent on PT. By integrating the variables at the two time points, the final independent interaction variables were formed. Table 2 provides an overview of these variables.

Covariates

In general, school achievement is influenced by a wide variety of individual factors, including prior knowledge, educational expectations, and family characteristics. The covariates controlled in this study are described in Table 2.

Descriptive statistics and correlation analysis

Table 3 presents the descriptive statistics for the variables in the analysis. In general, more than half of the sample students received PT (with significant intensity or frequency) during the summer vacation and school semester, whereas 30% of the students did not attend PT at either time point. As for the PT contents, 38.9% of the students previewed in-class knowledge at the two time points, 36.3% reviewed the knowledge during the semester and previewed the knowledge during the summer vacation, and 24.4% reviewed the knowledge in class at both time points. In terms of the PT form of instructional organization, more students participated in-class tutoring, followed by those who participated in one-to-one tutoring and online tutoring (the least number of students). Regarding the qualifications of the tutors, the majority were regular school teachers and cram school teachers. However, it should be noted that some students might be taught by more than one tutor, some might participate in both one-to-one and online PT, and some might learn more than one type of instructional content at a time (e.g., previewing and reviewing in-class knowledge during the summer vacation). Thus, the sum of the percentage might be more than 100%.

As for the variable of educational expectations, 7% hoped to obtain a senior high school diploma (or lower), 40.7% hoped to obtain a junior college or undergraduate degree, 52% hoped to obtain a master's degree, and

Table 3 Descriptive statistics of the variables used for analysis

Name of variable	N	%	Name of variable	N	%
PT participation			PT teacher qualification		
PT–(semester)[a] none–(summer vacation) none	600	30.2	University student–university student	161	8.1
PT–none–occasional	18	0.9	University student–not university student	53	2.7
PT–none–regular	105	5.3	Not university student–university student	50	2.5
PT–occasional–none	6	0.3	Regular school teacher–regular school teacher	311	15.7
PT–occasional–occasional	27	1.4	Regular school teacher–not regular school teacher	80	4.0
PT–occasional–regular	21	1.1	Not regular school teacher–regular school teacher	85	4.3
PT–regular–none	101	5.1	University teacher–university teacher	93	4.7
PT–regular–occasional	35	1.8	University teacher–not university teacher	32	1.6
PT–regular–regular	1075	54.1	Not university teacher–university teacher	34	1.7
PT instruction content			Cram school-cram school	415	20.9
Preview–preview	773	38.9	Cram school–not cram school	90	4.5
Preview–review	404	20.3	Not cram school-cram school	110	5.5
Preview–others	190	9.6	Female	962	48.4
Review–preview	721	36.3	Educational expectations		
Review–review	486	24.4	Senior high school or less	140	7.0
Review–others	171	8.6	Junior college, undergraduate	809	40.7
Others–preview	251	12.6	Master's	440	22.1
Others–review	146	7.3	PhD	599	30.1
Others–others	237	11.9	Mother's education		
PT form of instructional organization			Primary school or less	97	4.9
One-to-one–one-to-one	229	11.5	Junior high school	502	25.3
One-to-one–not one-to-one	113	5.7	Technical secondary school, vocational high School, senior school	786	39.5
Not one-to-one–one-to-one	57	2.9	Junior college, undergraduate	542	27.3
Class–class	854	43.0	Master's or PhD.	61	3.1
Class–not class	100	5.0	Financial condition of family		
Not class–class	154	7.7	Very poor	49	2.5
Online–online	99	5.0	Somewhat poor	96	4.8
Online–not online	59	3.0	Moderate	1578	79.4
Not online–online	41	2.1	Somewhat rich	234	11.8
Variables (number of valid cases)	Mean	S.D.	Very rich	31	1.6
Hours spent on PT	51.652	40.454			
Prior achievement	80.559	19.291	Total case no.	1988	100

[a]The first index is for the characteristic (status) of school semester PT, while the second one is for that of summer vacation PT

approximately one-third hoped to obtain a doctorate degree. Regarding the education of the participants' mothers, 4.9% were graduates from a primary (or lower) school, 25.3% had junior high school education, and 39.5% were technical secondary school, vocational high school or senior school graduates. These statistics indicated that the educational level of the participants' mothers was not very high. As for the students' family backgrounds, most of the families were perceived by the students as being at the middle economic level. Overall, the average of the students' prior scores was approximately 80.6. Regarding the inter-correlations analysis, the findings are shown in Appendix A (Table 7).

Methods

In order to explore the impact of PT on mathematics achievement and address the research questions, two regression models were applied. First, this study used a hierarchical linear model to analyze the association between PT participation with academic achievement. Second, unconditional quantile regression was used to explore the heterogeneity of the relationship between PT participation and academic achievement in different distributions of the latter. Moreover, the hierarchical linear model was used to explore the correlations between PT instruction contents, PT forms of instructional organization, hours

spent on PT, PT teacher qualifications, and academic achievement.

Hierarchical linear model

According to the characteristics of the students, this study used a two-level (student and class level) hierarchical linear model (hereafter referred to as "HLM") to examine the association between PT and academic achievement. When conducting the HLM analysis, the null model estimates were produced without any characteristics of the students and classes. More specifically, each null model estimated the variance within and between the classes for the corresponding student achievement. This variance within and between classes is further used to calculate the intraclass correlation coefficient (hereafter referred to as "ICC"), which measures the proportion of the variance of the student outcome (at the class level) to the total variance. In this case, if the ICC is greater than 0.059 (Cohen 1988), then a hierarchical linear model analysis is necessary. Level-1 model (students) and Level-2 model (classes)

$$y_{ij} = \beta_{0j} + \beta_{1j}PT_{ij} + \beta_{2j}X_{ij} + \epsilon_{ij} \tag{1}$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{2}$$

where y_{ij} is the grade of i from class j, β_{0j} is the average grade of class j students, PT_{ij} is the independent variable associated with PT, β_{1j} is the average coefficient of the independent variable, X_{ij} is the covariate, β_{2j} is the average coefficient of the covariate, ϵ_{ij} is the variation of i students of class j from the average grade of class j, γ_{00} is the average grade of all of the students, and u_{0j} is the variation in each class mean from the mean of the entire population.

Unconditional quantile regression

According to this study's hypothesis, the influence of mathematics PT on students' performance may vary for different groups of students with different achievement levels. To date, the majority of the studies on the relationship between PT participation and academic achievement (in different distributions of the latter) have used conditional quantile regression, as developed by Koenker and Bassett (1978). In this regard, conditional quantile regression only examines the heterogeneous effect of students' PT participation, with the same observed characteristics regarding their scores.

However, the present study is more interested in the impact of PT on academic achievement using the entire sample of heterogeneous workers (unconditional effects), rather than a sub-sample of students with specific characteristics (conditional effects). Thus, it adopted the unconditional quantile regression method to explore the associations between PT and students' academic achievement in different grades, as developed by Firpo et al. (2009). Within this framework, the re-centered influence function (hereafter referred to as "RIF") is an estimated pre-regression that serves as the outcome variable in a least-squares regression. The RIF is calculated as follows:

$$(Y; q_\tau, F_y) = q_\tau + \left[\frac{\tau - 1\{Y \le q_\tau\}}{f_y(q_\tau)} \right]$$

where Y is the outcome variable (student achievement), τ is the specific quantile, q_τ is the value of the outcome variable at this quantile, F_y is the cumulative distribution function of Y, $F_y(q_\tau)$ is the density of Y at q_τ, and $1\{Y \le q_\tau\}$ is the dummy variable indicating whether the outcome variable is below q_τ. The RIF is then included into the model as a dependent variable in the least-squares estimate. It should also be noted that class dummies were included in the model and the standard error of clustering (at the class level) was used to obtain more robust estimates.

Results

Overall effect of PT

Table 4 presents the results from the two-level HLM analysis. In order to provide a comprehensive description of the students' PT participation at the two time points, eight interaction dummy variables were reorganized. Specific to the fixed effects, the students who attended PT occasionally during the summer vacation and school semester experienced a significant negative effect on their performance, with a score that was a 0.239 standard deviation lower than the non-tutored students. However, it should be noted that the number of students who participated in occasional PT at both time points was relatively small, with only 27 students in the sample. Moreover, although there was no significant relationship between the other types of PT and the students' academic achievement, it should be emphasized that the effect of attending regular PT at the

Table 4 Results of the two-level HLM analysis for the effectiveness of PT

Fixed effect	Estimate	Std. error
Private tutoring		
PT–(semester) none–(summer vacation) none (reference group)		
PT–none–occasional	− 0.169	0.114
PT–none–regular	− 0.005	0.051
PT–occasional–none	0.118	0.194
PT–occasional–occasional	− 0.239*	0.093
PT–occasional–regular	0.024	0.105
PT–regular–none	0.007	0.051
PT–regular–occasional	− 0.070	0.082
PT–regular–regular	0.042	0.026
Prior knowledge	0.041**	0.000
Female	0.044*	0.022
Educational expectations		
Senior high school or lower (reference group)		
Junior college, undergraduate	0.071	0.046
Master's	0.203 **	0.052
PhD	0.229 **	0.050
Mother's education		
Primary school or less	0.057	0.053
Junior high school (reference group)		
Technical secondary school, vocational high school, senior school	− 0.031	0.028
Junior college, undergraduate	0.011	0.032
Master's or PhD	0.039	0.068
Family economy		
Very poor (reference group)		
Somewhat poor	− 0.412	0.084
Moderate	− 0.013	0.070
Somewhat rich	0.010	0.076
Very rich	− 0.093	0.110
Random effect		**Variance**
Level-1 variance		0.219
Level-2 variance		0.022
N		1988

Significance is denoted by *$p < 0.05$ and **$p < 0.01$

two time points on academic performance was 0.042, with a *p* value of 0.103.[1]

---

[1] Given the concerns about endogeneity and instrumental variables analysis, an instrumental analysis was conducted that used two rubrics for the survey items: "My classmates or friends participate in PT, so I choose to participate in PT"; and "Few classmates or friends participate in PT, so I choose not to participate in PT." The results (not reported here, but available upon request) showed that the findings of the instrumental variable method were roughly the same as those of the hierarchical linear model.

Heterogeneous effect on different groups of students

This study further investigated the effectiveness of regular PT across students with different mathematics achievement. In order to simplify the complex combination of characteristics, the following analysis focused on a sub-sample (1675) in which the students either received PT regularly at both time points or did not receive PT at either time point.

Table 5 reports the unconditional quantile regression results of the students' scores. The associations between PT and the students' academic achievement were obviously different in the different quantiles. For example, according to

Table 5 Unconditional quantile regression of the students' academic achievement

Variable	5th	10th	25th	50th	75th	90th	95th
PT–regular–regular	0.312**	0.416**	0.097	− 0.121	− 0.049	− 0.019	0.003
	(0.094)	(0.123)	(0.078)	(0.063)	(0.034)	(0.029)	(0.029)
Intercept	− 6.796**	− 9.161**	− 7.639**	− 2.488**	− 0.109	0.595**	0.798**
	(0.533)	(0.658)	(0.345)	(0.330)	(0.168)	(0.109)	(0.108)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1675	1675	1675	1675	1675	1675	1675

The covariates in this model are the same as those in the HLM model. All unconditional quantile regressions are estimated with cluster–robust standard errors. Significance is denoted by *$p < 0.05$ and **$p < 0.01$

the results from the 5th to 95th quantiles, regular PT at both time points had a positive significant impact on the students at the bottom of the achievement distribution. This indicates that regular PT at both time points benefits students in the lower quantile. Conversely, in the middle of the distribution, regular PT at both time points had a significant negative impact on the students' achievement. This suggests that PT can be harmful to students at the middle of the achievement distribution. Regular PT also had a negative impact on the students' scores at the top of the achievement distribution. However, the impact was not as significant as that for the students at the middle of the achievement distribution.

Let us now consider the unconditional quantile regression results in more detail, for the purpose of describing how these estimated achievement differences evolved across the entire achievement distribution, rather than simply focusing on the selected quantiles presented in Table 5. Figure 1 reports the changes in the regression coefficient of regular PT throughout the school semester and summer vacation (at different quantiles), after controlling for the covariates. According to the findings, the effect was statistically significant for the students in the lower quantile. In general, the influence of regular PT (at both time points) on the scores is largest and reaches a significant level at the low quantile. Meanwhile, as the number of the quantile increases, the effect value decreases and drops to the lowest value (between the 50th and 60th quantiles), reaching an insignificant negative direction. Additionally, as the number of the quantile increases, the effect continues in an insignificant negative direction, only becoming greater than 0 around the 95th quantile.

Detailed analysis of the effect: presage variables, process variables, and intensity

The following analysis uses the same sub-sample as the previous section. Among the regular tutored and non-tutored students at both time points, the effects of the contents, forms of instruction organization, strength, and qualifications of tutors were further analyzed, as shown in Table 6.

As for the presage variables, one-to-one tutoring had a significant negative impact on the students' performance. More specifically, the students who participated in one-to-one PT at both time points performed 0.13 standard deviation lower than their counterparts, while online tutoring during the school semester and summer vacation had a significant negative effect (up to 0.205) on the students' performance.

Regarding the process variables, the scores of the students who previewed the contents through PT at both time points was 0.077 standard deviation higher than those of the students who did not participate in PT, with a marginal significance ($p = 0.054$). Moreover, at both time points, using PT to review the contents had a marginally significant negative impact on the students' performance ($p = 0.052$). As for PT intensity, the findings showed that there was no significant effect. Meanwhile, the effects of squared PT intensity (excluded from the final model) were also not significant.
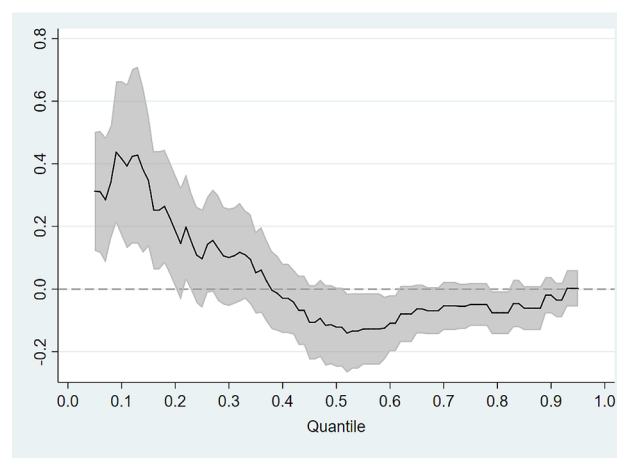

Fig. 1 The changes in the regression coefficient of regular PT. Note: This graph was plotted using the outcomes of the unconditional quantile regressions for τ between 0.05 and 0.95. The solid line indicates the unconditional quantile regression results. The coefficient intervals are shaded and constructed at the 95% level to show the significance of the estimates

Table 6 Results of the two-level HLM analysis for the effects of the contents, instructional organization, and qualifications of tutors, etc

Fixed effect	Estimate	Std. error
Instructional content		
PT–(semester) none–(vacation) none (reference group)		
Preview–preview	0.077	0.040
Preview–review	− 0.013	0.047
Preview–others	− 0.025	0.063
Review–preview	0.005	0.036
Review–review	− 0.085	0.044
Review–others	0.024	0.065
Others–preview	− 0.060	0.052
Others–review	0.015	0.064
Others–others	0.108*	0.055
Instructional organization		
None–none (reference group)		
One-to-one–one-to-one	− 0.130*	0.055
One-to-one–not one-to-one	0.003	0.080
Not one-to-one–one-to-one	− 0.006	0.092
Class–class	− 0.071	0.053
Class–not class	− 0.147	0.111
Not class–class	− 0.009	0.085
Online–online	− 0.205**	0.061
Online–not online	0.045	0.084
Not online–online	− 0.037	0.092
Hours spent on PT		
Hours	0.001	0.001
Hours squared	–	–
Qualifications of tutors		
None–none (reference group)		
University student–university student	0.033	0.047
University student–not university student	0.035	0.097
Not university student–university student	0.067	0.089
Regular school teacher–regular school teacher	0.066	0.039
Regular school teacher–not regular school teacher	− 0.033	0.072
Not regular school teacher–regular school teacher	0.245**	0.073
University teacher–university teacher	− 0.043	0.056
University teacher–not university teacher	− 0.207*	0.103
Not university teacher–university teacher	− 0.351**	0.127
Cram school-cram school	0.092**	0.037
Cram school–not cram school	− 0.009	0.063
Not cram school-cram school	0.062	0.072
Prior knowledge	0.040**	0.001
Female	0.027	0.024
Educational expectations		
Senior high school or less (reference group)		
Junior college, undergraduate	0.096	0.051
Master's	0.238**	0.057
PhD	0.279**	0.055
Mother's education		
Primary school or less	0.080	0.058
Junior high school (reference group)		

Table 6 (continued)

Fixed effect	Estimate	Std. error
Technical secondary school, vocational high school, senior school	− 0.010	0.030
Junior college, undergraduate	0.023	0.036
Master's or PhD	0.077	0.073
Family economy		
Very poor (reference group)		
Somewhat poor	− 0.074	0.094
Moderate	− 0.051	0.078
Somewhat rich	− 0.052	0.084
Very rich	− 0.098**	0.124

Random effect		Variance
Level-1		0.215
Level-2		0.022
N		1675

Significance is denoted by *$p < 0.05$ and **$p < 0.01$

The teachers from cram schools at both time points had a significant positive impact on the students' performance. Additionally, the university student tutors potentially negatively impacted the students' achievement, whereas the regular school teacher tutors potentially positively impacted the students' achievement.

Discussion

The present study conducted a comprehensive analysis of the effectiveness of PT on students' academic achievement. Moreover, the results regarding the effectiveness of different forms of PT and how their effects are heterogeneous are entirely new to the literature.

Overall effect of PT

Following the trends of previous studies, the first step in assessing the quality of PT was to measure the overall effect of PT on the students' academic achievement, even though it might not have provided comprehensive information on the forms of PT for theory and practice with consideration of the form of PT (Bray 2014). However, the results of the present study indicated that participating in regular PT throughout the summer vacation and school semester only had a weak, marginally significant positive effect ($p = 0.103$) on the students' mathematics achievement. This finding is in line with previous studies that applied similar longitudinal designs (e.g., Kuan 2011; Lee 2013; Zhang 2013; Byun 2014; Guill et al. 2019; Ömeroğulları et al. 2020), which did not show a significant positive effect of effect size, thus contradicting the students' expectations.

The present results also enriched previous studies by distinguishing PT at two time points. For example, occasional PT at both time points can negatively impact mathematics achievement. However, the results might be biased, since the sample size was relatively small. In addition, these results imply that instructional time should be significantly increased and frequency should be maintained over a longer period of time to see small (insignificant) improvements in mathematics learning. In other words, substantially increased instructional time might not benefit students, which is consistent with Carroll's model. To provide a more comprehensive analysis of PT effectiveness, the quality of PT should be evaluated to determine what forms of PT can help students achieve greater improvement in mathematics learning, or whether some groups of students can benefit more from PT.

Heterogeneous effect of PT on different groups of students

Since the aforementioned model did not find a significant positive effect of PT on the students' academic achievement, more detailed data were necessary. Thus, unconditional quantile regression was applied to examine the heterogeneous effect across students with different mathematics achievement (Zhang 2013), especially the sub-samples of students who regularly received PT during the summer vacation and following semester, and those who did not receive any PT throughout the summer vacation and the semester.

In contrast to the results of existing studies (Choi and Park 2016; Zhang 2013), the effect size of the present study was positive and significant in the lower quantiles, whereas it was negative and insignificant in the middle and upper

quantiles. However, limited studies have discussed the reasons for this heterogeneous effect. Possible explanations include the different learning styles of the students from different groups, the ceiling effect of the testing tool, and the quality of PT instruction for different groups of students (e.g., coherence between PT instruction and the learning needs of students) and future studies need to address this issue using a different methodology, such as a qualitative).

In the present study, the students in the lower quantiles enhanced their ability to solve routine problems by regularly receiving PT instruction that extended their learning time and provided examination skills training (Yung 2020). Meanwhile, the students in the middle and top quantiles did not necessarily benefit from PT and in fact may even be negatively impacted by PT (especially for the middle-class students). These students might need substantial help with more complex mathematical ideas and the ability to solve non-routine problems with higher cognitive demand (Stein, Grover, and Henningsen 1996) to see improvement in their mathematics achievement. However, the latter may not be easily enhanced by merely receiving PT to extend learning time, especially when the PT is of low instructional quality. Moreover, regular PT might conflict with students' formal school learning (e.g., occupying time better spent on homework), especially when PT is not designed to provide students with the opportunity to learn content of higher cognitive demand that perhaps should be based on inquiry learning (Yackel and Cobb 1996) rather than on the examination skills PT typically focuses on (Yung 2020).

Detailed analysis of the effectiveness of PT

The main contribution of the present study was to provide a detailed analysis on the quality of PT instruction and provide both statistical evidence and insights for future deeper analysis of each aspect of the quality of PT instruction. Several presage variables, process variables, and an intensity variable were also discussed, in addition to the characteristics of the sub-sample in which the students received regular PT at both time points or received PT at neither time point.

For the presage variables, the results for the qualifications of tutors were generally consistent with experience, but inconsistent with previous results (e.g., Ömeroğulları et al. 2020). This indicated that both regular school and cram school teachers benefited students' mathematics learning, whereas university teacher/student tutors did not have such an effect. In this regard, the PT instruction might have had a specific pedagogy that differed from that of regular school instruction (Wang et al. 2018, 2019) such as one-to-one teaching, diagnostic analysis of the students' mathematics learning (Herppich et al. 2013) or special examination preparation skills (Yung 2020). Thus, more research on teaching expertise and teacher development is necessary such as comparing the PT instruction of regular school teachers with

that of cram school teachers or comparing the professional development designed by a cram school with that designed by a regular school. These results can be used to better identify the qualifications of tutors, rather than simply using the classifications in existing studies. Additionally, more indices on the qualities/qualifications of tutors should be developed to complement the existing, overly simplified qualifications for tutors that are merely based on their backgrounds.

Regarding the results of instructional organization, the students were negatively impacted from online PT. This finding is consistent with Byun (2014), which found a non-significant negative effect. This also implies that online instruction requires more empirical evidence in the light of its rapid expansion. Contrary to the expectation, though one-to-one tutoring consumed more resources, it impacted mathematics achievement negatively. Though previous studies did not find a positive effect of one-to-one tutoring, Zhang and Liu (2016) indicated that large class-size tutoring was effective for improving students' analytical and problem-solving skills for examinations. However, these results require more exploration. One possibility is that the current model and data could not avoid all of the self-selection bias in PT participation (Ömeroğulları et al. 2020). Another direction might be to conduct a deeper analysis of one-to-one tutoring, e.g., determining if tutors with diverse backgrounds need specific expertise to be successful in such tutoring.

As for the process variables, the PT contents variable was entirely new to the literature on the effectiveness of different types of PT, which was a limitation in previous studies (Zhang et al. 2020). Compared to the review of instruction contents, the regular preview of new content in school instruction that continued throughout the summer vacation and the school semester helped the students obtain higher mathematics achievement at the end of the school semester. Although there were some contrasting points of view about previewing, such as reducing the students' interest/curiosity and disturbing the teachers' instruction in class (since some of the students had already previewed the contents) (Liang 2014; Wang et al. 2018, 2019), the evidence supports the effectiveness of previewing school contents through PT.

At the middle school level, mathematics learning generally focuses on mathematics knowledge and routine problem-solving skills, such as calculating mathematical expressions and solving equations, with testing at the end of the semester that determines whether low cognitive level/demand (Stein et al. 1996) students can benefit from previewing by practicing. In this regard, the results of the present study indicated that lower quantile students can benefit from PT, since previewing helped them enhance their routine problem-solving abilities, which in turn, improved their mathematics achievement. However, it should be cautioned that these results might not be applicable in mathematics learning that requires high cognitive levels or at the high school level, when the learning contents focus on

a deeper understanding of mathematical concepts and more complex (non-routine) problem-solving skills (Yackel and Cobb 1996). Thus, more studies should focus on the effect of PT on students' mathematics learning at high cognitive levels.

Meanwhile, following Guill et al. (2019), more process variables related to instructional practices should be designed, such as structure, challenge, and support in PT, although no significance was found in their study. Moreover, the structure of presage and process variables should be further explored, e.g., determining whether the process mediates the effect of presage variables on students' academic achievement.

For the review of the learning contents at school, a weak negative but marginally significant effect was found, which can be partly explained by the concept of instructional coherence (Chen and Li 2010). This concept is defined as causally linked activities/events, in terms of their instructional content and the meaningful discourse reflecting the connectedness of topics (Chen and Li 2010) (within a lesson or across lessons), which can improve students' mathematics learning through coherent and conceptual understanding (Fernandez et al. 1992). With respect to reviewing learned contents, instructional coherence might be a more complex issue, since the tutors require detailed knowledge (information) about what the students have learned, and they must carefully select the mathematics questions accordingly. This is particularly difficult for students from different classes/schools with different learning processes. Some studies have also shown that incoherent reviews might not only conflict with school teaching (Liang 2014), but they might also occupy the learning time after school (Liang 2014; Ömeroğulları et al. 2020). Alternatively, previewing is simply a glance at fundamental concepts, algorithms, and theorems that appear in the textbook, with which it is much easier to achieve a high level of coherence with school learning, and which is useful in enhancing students' achievement in routine problem-solving skills. Overall, since the results for the process variables of PT quality still somewhat vary, more comprehensive studies should analyze the pedagogy of PT and the different learning styles of students with different mathematics achievement (Chi et al. 2004; Chu et al. 2017; Wang et al. 2018, 2019).

Finally, for the intensity variable, when the students participated in regular PT during the summer vacation and school semester, there was no evidence that the increase in intensity enhanced their mathematics achievement, which was similar to Ömeroğulları et al. (2020) and different from Liu's (2012) cross-sectional study. This result suggests that the time spent on PT will not increase the total amount of effective learning time when it is not of high instructional quality. In addition, previous studies have shown that effective learning time tends to decrease when tutored students pay less attention during regular school lessons (Liang 2014; Ömeroğulları et al. 2020), which is consistent with the framework of Carroll's model. Both of these explanations imply that PT effectiveness decreases (or

remains constant) as the intensity of PT increases. Furthermore, no inflection point (or an inverse U curve) was found. This result is in contrast to some cross-sectional studies (e.g., Liu 2012; Wang and Guo 2017), which found an inverse U curve and supported the over-scheduling hypothesis in broader extracurricular activities (Fredricks 2012). Thus, this hypothesis requires further discussion in future research.

Suggestions for future research

Although the results in the present study supplement those of previous studies, through a detailed analysis of the PT effect, key "pieces of the puzzle" are still missing. Hence, more comprehensive studies should be designed to provide more evidence and support the aforementioned explanations. For example, further expansion of the theoretical model in the present study could include: the process variables of PT; the data that reflects the effectiveness of PT such as the non-cognitive aspects of learning outcomes (Guill et al. 2019); the heterogeneous effect on different aspects of academic achievement such as geometry or algebra achievement; and the heterogenous effect on students of different cognitive levels (Wang et al. 2018, 2019).

Conclusion and implications

The present study conducted a comprehensive analysis based on a specially designed retrospective longitudinal survey of private supplementary mathematics tutoring among middle school students in China. In conclusion, our findings do not support private tutoring as an definitely effective strategy to improve students' school mathematics achievement, even when regular PT continues throughout the summer vacation and the school semester. This study also found that under specific circumstances, some presage variables, such as tutor qualifications, time points, forms of instructional organization, and process variables (instructional contents and context variables), can explain the effectiveness of PT. In addition, the lower quantile students might benefit more from PT when compared to others.

Regarding the practical and policy implications, this study indicates that parents should carefully select PT for their children. It also suggests that the government should provide more comprehensive professional guidelines to regulate the PT industry, especially online instruction which becomes more and more popular.

Appendix A

See Table 7.

Table 7 Correlation analysis

	Variable	1	2	3	4	5	6	7	8
1	PT-(semester) none-(summer vacation) none	1.00							
2	PT-none-occasionally	**− 0.06**	1.00						
3	PT–none–regular	**− 0.16**	− 0.02	1.00					
4	PT-occasionally-none	− 0.04	− 0.01	− 0.01	1.00				
5	PT-occasionally-occasionally	**− 0.08**	− 0.01	− 0.03	− 0.01	1.00			
6	PT-occasionally-regular	**− 0.07**	− 0.01	− 0.02	− 0.01	− 0.01	1.00		
7	PT–regular–none	**− 0.15**	− 0.02	**− 0.05**	− 0.01	− 0.03	− 0.02	1.00	
8	PT-regular-occasionally	**− 0.09**	− 0.01	− 0.03	− 0.01	− 0.02	− 0.01	− 0.03	1.00
9	PT–regular–regular	**− 0.71**	**− 0.10**	**− 0.26**	**− 0.06**	**− 0.13**	**− 0.11**	**− 0.25**	**− 0.15**
10	Content: preview–preview	**− 0.52**	**− 0.08**	**− 0.19**	− 0.03	**0.05**	− 0.01	**− 0.17**	**0.07**
11	Content: preview–review	**− 0.33**	**− 0.05**	**− 0.12**	− 0.03	0.03	0.01	**− 0.11**	0.02
12	Content: preview–others	**− 0.21**	− 0.03	**− 0.08**	− 0.02	0.04	0.00	**− 0.06**	− 0.02
13	Content: review–preview	**− 0.49**	**− 0.07**	**− 0.18**	− 0.02	− 0.01	0.04	**− 0.15**	0.03
14	Content: review–review	**− 0.37**	**− 0.05**	**− 0.13**	− 0.03	**0.05**	0.03	**− 0.12**	0.01
15	Content: review–others	**− 0.20**	− 0.03	**− 0.07**	− 0.02	0.03	0.00	**− 0.05**	− 0.03
16	Content: others–preview	**− 0.25**	**− 0.04**	**− 0.09**	− 0.02	0.03	− 0.01	**− 0.07**	0.00
17	Content: others–review	**− 0.18**	− 0.03	**− 0.07**	− 0.02	0.03	0.03	− 0.04	− 0.02
18	Content: others–others	**− 0.24**	**− 0.04**	**− 0.09**	− 0.02	**0.09**	0.01	**− 0.06**	− 0.03
19	Organization: one-to-one–one-to-one	**− 0.23**	− 0.03	**− 0.09**	− 0.02	**0.05**	− 0.01	**− 0.08**	0.05
20	Organization: one-to-one -not one-to-one	**− 0.15**	**0.14**	**0.27**	− 0.01	− 0.03	0.04	**− 0.05**	0.02
21	Organization: not one-to-one - one-to-one	**− 0.11**	− 0.02	− 0.04	**0.05**	0.01	− 0.02	**0.22**	0.02
22	Organization: class–class	**− 0.57**	**− 0.08**	**− 0.20**	**− 0.05**	− 0.01	**0.06**	**− 0.20**	0.00
23	Organization: class–not class	**− 0.14**	**0.10**	**0.64**	− 0.01	− 0.03	0.02	**− 0.05**	0.00
24	Organization: not class–class	**− 0.19**	− 0.03	**− 0.07**	**0.12**	− 0.02	0.01	**0.60**	**0.08**
25	Organization: online–online	**− 0.15**	− 0.02	**− 0.05**	− 0.01	**0.15**	0.02	**− 0.05**	**0.09**
26	Organization: online–not online	**− 0.11**	**0.14**	**0.09**	− 0.01	0.03	0.01	− 0.03	**0.11**
27	Organization: not online–online	**− 0.10**	− 0.01	− 0.03	**0.06**	− 0.02	0.02	**0.11**	0.01
28	PT strength	**− 0.84**	**− 0.08**	0.04	**− 0.06**	**− 0.08**	0.02	**− 0.16**	− 0.04
29	Tutor qualifications: university student–university student	**− 0.20**	− 0.03	**− 0.07**	− 0.02	**0.09**	0.04	− 0.04	0.04
30	Tutor qualifications: university student–not university student	**− 0.10**	**0.05**	**0.18**	**0.05**	0.01	0.01	− 0.02	**0.05**
31	Tutor qualifications: not university student–university student	**− 0.10**	− 0.02	− 0.04	− 0.01	0.04	0.01	**0.12**	0.03
32	Tutor qualifications: regular school teacher–regular school teacher	**− 0.27**	**− 0.04**	**− 0.10**	− 0.02	0.03	0.00	**− 0.08**	0.02
33	Tutor qualifications: regular school teacher–not regular school teacher	**− 0.13**	− 0.02	**0.21**	0.04	− 0.02	**0.08**	− 0.02	− 0.03
34	Tutor qualifications: not regular school teacher–regular school teacher	**− 0.12**	− 0.02	− 0.04	**0.08**	− 0.02	− 0.02	**0.27**	**0.05**
35	Tutor qualifications: university teacher -university teacher	**− 0.15**	− 0.02	**− 0.05**	− 0.01	0.02	0.02	**− 0.05**	**0.06**
36	Tutor qualifications: university teacher -not university teacher	**− 0.08**	0.03	**0.18**	**0.07**	0.02	− 0.01	− 0.03	0.01
37	Tutor qualifications: not university teacher -university teacher	**− 0.07**	− 0.01	− 0.03	− 0.01	− 0.02	− 0.01	**0.09**	0.01
38	Tutor qualifications: cram school-cram school	**− 0.34**	**− 0.05**	**− 0.12**	− 0.03	− 0.03	0.01	**− 0.11**	0.02
39	Tutor qualifications: cram school–not cram school	**− 0.14**	0.03	**0.26**	0.03	0.00	− 0.02	**− 0.05**	0.01
40	Tutor qualifications: not cram school-cram school	**− 0.16**	− 0.02	**− 0.06**	**0.11**	− 0.03	0.02	**0.25**	0.00
41	Prior knowledge	**− 0.11**	− 0.03	0.00	0.03	− 0.02	0.04	− 0.01	− 0.01
42	Female	**− 0.16**	**0.05**	**0.05**	− 0.02	0.02	− 0.04	0.01	0.01
43	Family economy	**− 0.10**	− 0.02	0.02	− 0.04	− 0.01	− 0.01	− 0.02	0.01
44	Mother's education: primary school or less	**0.11**	− 0.02	− 0.01	0.03	− 0.01	− 0.02	0.02	− 0.01
45	Mother's education: junior high school	**0.08**	0.01	0.02	− 0.01	− 0.02	− 0.01	0.01	0.00
46	Mother's education: technical secondary school, vocational high school, senior school	0.01	0.01	− 0.02	− 0.03	0.03	0.00	0.02	**− 0.05**
47	Mother's education: junior college, undergraduate	**− 0.14**	0.00	0.01	0.03	− 0.02	0.04	− 0.03	**0.06**

Table 7 (continued)

	Variable	1	2	3	4	5	6	7	8
48	Mother's education: Master's or PhD.	0.00	− 0.02	0.00	− 0.01	0.03	− 0.02	− 0.03	0.02
49	Self-expectation of education: senior high school or less	**0.14**	− 0.01	− 0.04	− 0.02	0.04	− 0.01	0.02	0.02
50	Self-expectation of education: junior college, undergraduate	0.04	0.04	− 0.01	0.03	− 0.02	− 0.01	0.02	− 0.03
51	Self-expectation of education: Master's	**− 0.13**	0.00	0.03	0.01	0.01	0.03	0.01	0.02
52	Self-expectation of education: PhD.	− 0.01	− 0.04	0.00	− 0.04	− 0.01	− 0.01	− 0.04	0.00

	Variable	9	10	11	12	13	14	15	16
1	PT-none-none								
2	PT-none-occasionally								
3	PT–none–regular								
4	PT-occasionally-none								
5	PT-occasionally-occasionally								
6	PT-occasionally-regular								
7	PT–regular–none								
8	PT-regular-occasionally								
9	PT–regular–regular	1.00							
10	Content: preview–preview	**0.63**	1.00						
11	Content: preview–review	**0.40**	**0.58**	1.00					
12	Content: preview–others	**0.26**	**0.36**	**0.26**	1.00				
13	Content: review–preview	**0.60**	**0.53**	**0.47**	**0.20**	1.00			
14	Content: review–review	**0.44**	**0.33**	**0.69**	**0.18**	**0.57**	1.00		
15	Content: review–others	**0.24**	**0.22**	**0.25**	**0.71**	**0.30**	**0.32**	1.00	
16	Content: others–preview	**0.30**	**0.30**	**0.20**	**0.56**	**0.25**	**0.15**	**0.48**	1.00
17	Content: others–review	**0.21**	**0.19**	**0.37**	**0.43**	**0.20**	**0.36**	**0.54**	**0.63**
18	Content: others–others	**0.27**	**0.13**	**0.12**	**0.62**	**0.08**	**0.12**	**0.59**	**0.59**
19	Organization: one-to-one–one-to-one	**0.28**	**0.17**	**0.21**	**0.10**	**0.22**	**0.26**	**0.11**	0.07
20	Organization: one-to-one -not one-to-one	0.00	− 0.02	− 0.03	− 0.04	**0.07**	− 0.01	− 0.04	0.04
21	Organization: not one-to-one - one-to-one	0.02	0.02	**0.05**	0.02	0.03	0.04	0.01	0.02
22	Organization: class–class	**0.71**	**0.55**	**0.30**	**0.24**	**0.45**	**0.31**	**0.22**	**0.26**
23	Organization: class–not class	**− 0.15**	**− 0.11**	**− 0.06**	− 0.07	**− 0.09**	**− 0.07**	**− 0.06**	**− 0.06**
24	Organization: not class–class	**− 0.08**	**− 0.07**	**− 0.06**	− 0.04	0.01	**− 0.06**	− 0.04	0.00
25	Organization: online–online	**0.13**	**0.13**	**0.11**	**0.06**	**0.11**	**0.11**	0.06	0.05
26	Organization: online–not online	0.01	0.04	0.00	0.01	0.04	− 0.02	0.03	0.00
27	Organization: not online–online	**0.05**	0.01	0.01	**0.05**	0.04	0.03	**0.08**	**0.06**
28	PT strength	**0.87**	**0.58**	**0.38**	**0.27**	**0.58**	**0.42**	**0.25**	**0.31**
29	Tutor qualifications: university student–university student	**0.19**	**0.13**	**0.11**	**0.16**	**0.09**	**0.12**	**0.15**	**0.14**
30	Tutor qualifications: university student–not university student	− 0.02	0.01	0.02	− 0.02	− 0.02	0.01	− 0.02	0.02
31	Tutor qualifications: not university student–university student	0.04	**0.06**	**0.05**	0.04	0.04	**0.05**	0.03	0.02
32	Tutor qualifications: regular school teacher–regular school teacher	**0.33**	**0.28**	**0.22**	**0.10**	**0.27**	**0.18**	**0.11**	**0.11**
33	Tutor qualifications: regular school teacher–not regular school teacher	0.03	0.02	− 0.01	− 0.01	0.02	0.01	0.00	0.02
34	Tutor qualifications: not regular school teacher–regular school teacher	0.01	0.00	0.01	− 0.03	0.03	0.00	− 0.04	− 0.01
35	Tutor qualifications: university teacher -university teacher	**0.16**	**0.11**	**0.11**	**0.11**	**0.13**	**0.16**	**0.13**	**0.11**
36	Tutor qualifications: university teacher -not university teacher	− 0.01	− 0.02	− 0.01	− 0.03	− 0.03	− 0.04	− 0.01	− 0.01
37	Tutor qualifications: not university teacher -university teacher	0.04	0.03	0.03	**0.06**	0.01	0.04	0.03	0.02
38	Tutor qualifications: cram school-cram school	**0.43**	**0.34**	0.21	**0.16**	**0.33**	**0.23**	**0.17**	**0.12**
39	Tutor qualifications: cram school–not cram school	0.03	0.02	0.02	**0.07**	0.02	0.01	**0.08**	**0.06**
40	Tutor qualifications: not cram school-cram school	**0.06**	0.03	− 0.03	− 0.03	**0.06**	− 0.01	− 0.02	0.02
41	Prior knowledge	**0.11**	**0.16**	0.03	**0.11**	**0.10**	− 0.04	**0.06**	**0.08**

Table 7 (continued)

	Variable	9	10	11	12	13	14	15	16
42	Female	**0.11**	**0.12**	**0.06**	− 0.01	**0.14**	**0.07**	0.04	0.00
43	Family economy	**0.10**	**0.06**	**0.06**	**0.08**	**0.06**	**0.07**	**0.07**	**0.05**
44	Mother's education: primary school or less	**− 0.09**	**− 0.06**	**− 0.05**	− 0.03	**− 0.08**	**− 0.06**	− 0.04	**− 0.05**
45	Mother's education: junior high school	**− 0.08**	**− 0.09**	**− 0.05**	**− 0.07**	**− 0.06**	− 0.04	**− 0.05**	0.00
46	Mother's education: technical secondary school, vocational high school, senior school	0.00	− 0.02	− 0.03	− 0.04	− 0.01	− 0.02	− 0.03	**− 0.06**
47	Mother's education: junior college, undergraduate	**0.12**	**0.12**	**0.10**	**0.11**	**0.11**	**0.09**	**0.11**	**0.08**
48	Mother's education: Master's or PhD.	0.01	0.02	0.03	**0.05**	0.00	0.00	0.01	0.04
49	Self-expectation of education: senior high school or less	**− 0.13**	**− 0.11**	**− 0.07**	− 0.04	**− 0.12**	**− 0.05**	**− 0.05**	**− 0.07**
50	Self-expectation of education: junior college, undergraduate	− 0.04	**− 0.06**	0.01	**− 0.10**	− 0.02	0.02	**− 0.07**	**− 0.05**
51	Self-expectation of education: Master's	**0.08**	**0.07**	0.01	0.02	**0.08**	0.02	0.04	0.00
52	Self-expectation of education: PhD.	0.04	**0.06**	0.01	0.11	0.01	− 0.01	**0.06**	**0.09**

	Variable	17	18	19	20	21	22	23	24
17	Content: others–review	1.00							
18	Content: others–others	**0.51**	1.00						
19	Organization: one-to-one- one-to-one	**0.09**	**0.11**	1.00					
20	Organization: one-to-one -not one-to-one	0.00	− 0.03	**− 0.09**	1.00				
21	Organization: not one-to-one - one-to-one	0.02	0.02	**− 0.06**	− 0.04	1.00			
22	Organization: class–class	**0.20**	**0.24**	**− 0.16**	**− 0.13**	− 0.03	1.00		
23	Organization: class–not class	**− 0.05**	**− 0.06**	− 0.03	− 0.01	**0.13**	**− 0.20**	1.00	
24	Organization: not class–class	− 0.02	− 0.04	− 0.03	**0.29**	0.00	**− 0.25**	**− 0.07**	1.00
25	Organization: online–online	0.05	**0.08**	0.04	− 0.03	0.03	**− 0.06**	− 0.01	0.01
26	Organization: online–not online	0.00	0.02	0.02	0.01	**0.08**	0.02	**0.05**	0.04
27	Organization: not online–online	0.03	**0.07**	0.03	**0.09**	0.04	0.01	0.00	0.04
28	PT strength	**0.22**	**0.26**	**0.31**	**0.13**	0.04	**0.62**	0.04	0.00
29	Tutor qualifications: university student–university student	**0.11**	**0.23**	**0.19**	0.03	**0.06**	**0.11**	0.01	0.04
30	Tutor qualifications: university student–not university student	0.01	− 0.03	0.00	**0.19**	**0.07**	− 0.03	**0.08**	0.02
31	Tutor qualifications: not university student–university student	0.04	0.02	**0.07**	− 0.01	**0.11**	0.02	− 0.02	**0.10**
32	Tutor qualifications: regular school teacher–regular school teacher	**0.11**	**0.07**	**0.17**	0.01	− 0.03	**0.29**	**− 0.05**	**− 0.05**
33	Tutor qualifications: regular school teacher–not regular school teacher	− 0.01	0.01	0.03	**0.05**	**0.07**	0.02	**0.20**	**− 0.05**
34	Tutor qualifications: not regular school teacher–regular school teacher	0.00	**− 0.05**	0.00	0.03	**0.10**	− 0.02	0.01	**0.23**
35	Tutor qualifications: university teacher -university teacher	**0.11**	**0.13**	**0.13**	− 0.01	0.00	**0.11**	− 0.03	− 0.02
36	Tutor qualifications: university teacher -not university teacher	− 0.02	0.01	− 0.01	**0.07**	**0.05**	− 0.02	**0.15**	0.01
37	Tutor qualifications: not university teacher -university teacher	0.04	0.02	0.04	0.00	**0.05**	0.03	− 0.03	**0.11**
38	Tutor qualifications: cram school-cram school	**0.12**	**0.15**	**0.06**	− 0.03	0.00	**0.41**	**− 0.10**	**− 0.07**
39	Tutor qualifications: cram school–not cram school	**0.11**	0.03	0.00	**0.10**	0.04	0.04	**0.25**	− 0.03
40	Tutor qualifications: not cram school-cram school	− 0.02	− 0.03	0.02	**0.07**	**0.08**	− 0.02	− 0.03	**0.27**
41	Prior knowledge	0.01	**0.06**	**− 0.12**	− 0.04	− 0.03	**0.18**	− 0.02	− 0.01
42	Female	0.02	− 0.03	**0.05**	0.03	0.01	**0.08**	**0.06**	0.02
43	Family economy	**0.08**	**0.09**	**0.09**	− 0.02	0.03	**0.07**	0.00	0.00
44	Mother's education: primary school or less	− 0.04	**− 0.05**	− 0.02	− 0.01	− 0.01	**− 0.09**	0.00	0.01
45	Mother's education: junior high school	− 0.01	**− 0.05**	**− 0.06**	− 0.02	− 0.03	**− 0.06**	0.01	0.01
46	Mother's education: technical secondary school, vocational high school, senior school	**− 0.05**	**− 0.08**	0.00	− 0.01	0.00	− 0.01	− 0.01	− 0.01
47	Mother's education: junior college, undergraduate	**0.08**	**0.12**	**0.06**	0.03	0.02	**0.12**	0.00	0.00
48	Mother's education: Master's or PhD.	− 0.01	**0.09**	**0.05**	0.01	0.04	− 0.01	0.00	− 0.01
49	Self-expectation of education: senior high school or less	**− 0.05**	**− 0.05**	0.02	− 0.03	**0.05**	**− 0.10**	0.00	− 0.01
50	Self-expectation of education: junior college, undergraduate	− 0.01	**− 0.08**	0.02	0.03	0.00	**− 0.08**	− 0.01	**0.05**
51	Self-expectation of education: Master's	− 0.01	0.01	− 0.03	0.02	0.02	**0.08**	0.03	0.02

Y. Zhang et al.

Table 7 (continued)

	Variable	17	18	19	20	21	22	23	24
52	Self-expectation of education: PhD.	0.04	**0.11**	−0.01	−0.03	−0.04	**0.06**	−0.02	**−0.06**

	Variable	25	26	27	28	29	30	31	32
17	Content: others–review								
18	Content: others–others								
19	Organization: one-to-one- one-to-one								
20	Organization: one-to-one -not one-to-one								
21	Organization: not one-to-one - one-to-one								
22	Organization: class–class								
23	Organization: class–not class								
24	Organization: not class–class								
25	Organization: online–online	1.00							
26	Organization: online–not online	−0.04	1.00						
27	Organization: not online–online	−0.03	−0.03	1.00					
28	PT strength	**0.15**	**0.08**	**0.06**	1.00				
29	Tutor qualifications: university student–university student	**0.14**	**0.06**	**0.07**	**0.21**	1.00			
30	Tutor qualifications: university student–not university student	0.02	**0.06**	0.02	**0.06**	**−0.05**	1.00		
31	Tutor qualifications: not university student–university student	−0.02	0.01	0.04	0.04	**−0.05**	−0.03	1.00	
32	Tutor qualifications: regular school teacher–regular school teacher	−0.03	0.02	0.00	**0.33**	0.02	−0.02	−0.04	1.00
33	Tutor qualifications: regular school teacher–not regular school teacher	0.04	−0.01	0.01	**0.10**	−0.01	**0.05**	**0.05**	**−0.09**
34	Tutor qualifications: not regular school teacher–regular school teacher	−0.03	0.02	0.00	0.02	−0.01	0.03	−0.02	**−0.09**
35	Tutor qualifications: university teacher -university teacher	**0.14**	**0.06**	**0.05**	**0.18**	**0.10**	0.02	**0.06**	0.04
36	Tutor qualifications: university teacher -not university teacher	−0.01	**0.07**	0.04	**0.05**	−0.01	**0.05**	**0.13**	−0.02
37	Tutor qualifications: not university teacher -university teacher	**0.09**	0.02	0.01	**0.05**	**0.07**	**0.17**	0.00	−0.01
38	Tutor qualifications: cram school-cram school	**0.14**	0.03	0.04	**0.39**	−0.01	**−0.05**	0.00	**−0.11**
39	Tutor qualifications: cram school–not cram school	0.01	**0.05**	−0.03	**0.10**	−0.01	0.02	**0.12**	**−0.07**
40	Tutor qualifications: not cram school-cram school	0.03	−0.03	**0.12**	**0.07**	−0.04	**0.11**	−0.01	−0.04
41	Prior knowledge	0.00	0.03	0.00	**0.10**	**−0.07**	0.00	**−0.06**	0.05
42	Female	0.01	0.02	**0.05**	**0.13**	−0.01	0.00	0.01	**0.07**
43	Family economy	0.03	0.01	−0.03	**0.12**	0.02	0.03	**0.08**	**0.07**
44	Mother's education: primary school or less	−0.02	−0.03	0.03	**−0.11**	−0.02	0.01	−0.01	**−0.08**
45	Mother's education: junior high school	**−0.05**	0.02	0.02	**−0.08**	−0.01	−0.02	−0.01	−0.04
46	Mother's education: technical secondary school, vocational high school, senior school	0.01	−0.01	−0.01	−0.02	−0.02	−0.01	−0.01	−0.02
47	Mother's education: junior college, undergraduate	0.04	−0.01	−0.03	**0.15**	0.00	0.02	0.02	**0.09**
48	Mother's education: Master's or PhD.	0.01	0.04	0.02	0.01	**0.09**	0.02	0.03	0.01
49	Self-expectation of education: senior high school or less	−0.04	−0.02	−0.01	**−0.14**	−0.01	−0.02	0.02	**−0.05**
50	Self-expectation of education: junior college, undergraduate	−0.02	−0.01	−0.01	−0.04	−0.03	0.04	0.02	**−0.05**
51	Self-expectation of education: Master's	0.01	0.01	0.00	**0.10**	0.03	0.00	0.01	**0.08**
52	Self-expectation of education: PhD.	0.04	0.01	0.02	0.04	0.01	−0.03	−0.04	0.00

	Variable	33	34	35	36	37	38	39	40
33	Tutor qualifications: regular school teacher–not regular school teacher	1.00							
34	Tutor qualifications: not regular school teacher–regular school teacher	−0.04	1.00						
35	Tutor qualifications: university teacher -university teacher	−0.02	**−0.05**	1.00					
36	Tutor qualifications: university teacher—not university teacher	0.01	**0.07**	−0.03	1.00				
37	Tutor qualifications: not university teacher—university teacher	**0.07**	**0.05**	−0.03	−0.02	1.00			
38	Tutor qualifications: cram school-cram school	**−0.05**	**−0.06**	0.03	−0.03	−0.02	1.00		
39	Tutor qualifications: cram school–not cram school	0.02	**0.10**	0.00	**0.05**	**0.06**	**−0.11**	1.00	
40	Tutor qualifications: not cram school-cram school	**0.12**	−0.01	−0.03	0.02	0.02	**−0.12**	**−0.05**	1.00
41	Prior knowledge	0.02	−0.01	**−0.09**	**−0.08**	−0.04	**0.08**	0.01	0.02

Table 7 (continued)

	Variable	33	34	35	36	37	38	39	40
42	Female	0.00	0.04	− 0.02	− 0.04	− 0.01	**0.11**	**0.06**	− 0.01
43	Family economy	0.00	0.01	**0.07**	0.03	**0.06**	**0.05**	**0.05**	− 0.01
44	Mother's education: primary school or less	0.01	0.04	− 0.02	− 0.01	0.02	− 0.04	− 0.02	0.01
45	Mother's education: junior high school	− 0.04	0.00	− 0.02	0.04	0.03	**− 0.06**	0.01	− 0.02
46	Mother's education: technical secondary school, vocational high school, senior school	− 0.01	0.00	− 0.01	− 0.01	− 0.04	− 0.04	0.01	0.00
47	Mother's education: junior college, undergraduate	0.04	− 0.01	0.02	− 0.02	− 0.01	**0.12**	0.00	0.00
48	Mother's education: Master's or PhD.	0.01	− 0.04	0.04	0.00	0.04	− 0.01	− 0.02	0.03
49	Self-expectation of education: senior high school or less	− 0.04	− 0.03	0.00	0.01	− 0.01	**− 0.05**	− 0.03	**− 0.06**
50	Self-expectation of education: junior college, undergraduate	− 0.01	0.04	− 0.01	0.01	0.00	**− 0.05**	0.03	− 0.01
51	Self-expectation of education: Master's	0.03	0.03	0.01	0.00	− 0.01	0.01	0.01	0.04
52	Self-expectation of education: PhD.	0.00	**− 0.05**	0.00	− 0.01	0.01	**0.08**	− 0.02	0.00

	Variable	41	42	43	44	45	46	47	48
33	Tutor qualifications: regular school teacher–not regular school teacher								
34	Tutor qualifications: not regular school teacher–regular school teacher								
35	Tutor qualifications: university teacher -university teacher								
36	Tutor qualifications: university teacher—not university teacher								
37	Tutor qualifications: not university teacher—university teacher								
38	Tutor qualifications: cram school-cram school								
39	Tutor qualifications: cram school–not cram school								
40	Tutor qualifications: not cram school-cram school								
41	Prior knowledge	1.00							
42	Female	**0.05**	1.00						
43	Family economy	0.02	**0.06**	1.00					
44	Mother's education: primary school or less	**− 0.09**	− 0.02	**− 0.10**	1.00				
45	Mother's education: junior high school	**− 0.14**	− 0.02	**− 0.10**	**− 0.13**	1.00			
46	Mother's education: technical secondary school, vocational high school, senior school	− 0.04	0.00	− 0.03	**− 0.18**	**− 0.47**	1.00		
47	Mother's education: junior college, undergraduate	**0.21**	0.05	**0.14**	**− 0.14**	**− 0.36**	**− 0.50**	1.00	
48	Mother's education: Master's or PhD.	0.04	− 0.04	**0.09**	− 0.04	**− 0.10**	**− 0.14**	**− 0.11**	1.00
49	Self-expectation of education: senior high school or less	**− 0.37**	**− 0.08**	**− 0.08**	**0.12**	**0.10**	0.00	**− 0.14**	− 0.04
50	Self-expectation of education: junior college, undergraduate	**− 0.18**	0.04	− 0.04	**0.05**	**0.14**	**0.06**	**− 0.19**	**− 0.09**
51	Self-expectation of education: Master's	**0.20**	0.05	**0.06**	**− 0.08**	**− 0.07**	− 0.01	**0.14**	**− 0.05**
52	Self-expectation of education: PhD.	**0.22**	− 0.04	0.03	**− 0.06**	**− 0.14**	**− 0.06**	**0.16**	**0.16**

	Variable	49	50	51	52
49	Self-expectation of education: senior high school or less	1			
50	Self-expectation of education: junior college, undergraduate	**− 0.23**	1		
51	Self-expectation of education: Master's	**− 0.15**	**− 0.44**	1	
52	Self-expectation of education: PhD.	**− 0.18**	**− 0.54**	**− 0.35**	1

Significant values are shown in bold ($p < .05$). $N = 1988$

References

Baker, D. P., Akiba, M., Letendre, G. K., & Wiseman, A. W. (2001). Worldwide shadow education: Outside-school learning, institutional quality of schooling, and cross–national mathematics achievement. *Educational Evaluation and Policy Analysis, 23*(1), 1–17.

Bray, M. (2009). *Confronting the shadow education system: What government policies for what private tutoring?*. Paris: UNESCO International Institute for Educational Planning.

Bray, M. (2014). The impact of shadow education on student academic achievement: Why the research is inconclusive and what can be done about it. *Asia Pacific Education Review, 15*(3), 381–389.

Bray, M., Kobakhidze, M. N., & Suter, L. E. (2020). The challenges of measuring outside–school–time educational activities: Experiences and lessons from the Programme for International Student Assessment (PISA). *Comparative Education Review*. https://doi.org/10.1086/706776.

Bray, M., & Lykins, C. (2012). *Shadow education: Private supplementary tutoring and its implications for policy makers in Asia*. Pok Fu Lam: Asian Development Bank Comparative Education Research Centre (CERC), Faculty of Education, The University of Hong Kong.

Byun, S. Y. (2014). Shadow education and academic success in South Korea. In H. Park & K. Kim (Eds.), *Korean education in changing economic and demographic contexts* (pp. 39–58). Dordrecht: Springer.

Byun, S. Y., Chung, H. J., & Baker, D. P. (2018). Global patterns of the use of shadow education: Student, family, and national influences. In H. Park & G. Kao (Eds.), *Research in the sociology of education* (Vol. 20, pp. 71–105). Bingley: Emerald Publishing Limited.

Campbell, R. J., Kyriakides, L., Muijs, R. D., & Robinson, W. (2003). Differential teacher effectiveness: Towards a model for research and teacher appraisal. *Oxford Review of Education, 29*(3), 347–362.

Carroll, J. B. (1963). A model of school learning. *Teacher College Record, 64,* 723–733.

Chappell, S., Nunnery, J., Pribesh, S., & Hager, J. (2011). A meta-analysis of supplemental educational services (SES) provider effects on student achievement. *Journal of Education for Students Placed at Risk, 16*(1), 1–23.

Chen, X., & Li, Y. (2010). Instructional coherence in Chinese mathematics classroom—A case study of lessons on fraction division. *International Journal of Science and Mathematics Education, 8*(4), 711–735.

Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition & Instruction, 22*(3), 363–387.

Choi, Y., & Park, H. (2016). Shadow education and educational inequality in South Korea: Examining effect heterogeneity of shadow education on middle school seniors' achievement test scores. *Research in Social Stratification and Mobility, 44,* 22–32.

Chu, H. C., Chen, J. M., & Tsai, C. L. (2017). Effects of an online formative peer-tutoring approach on students' learning behaviors, performance and cognitive load in mathematics. *Interactive Learning Environments, 25*(2), 203–219.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Laurence Erlbaum Associates.

Cole, R. (2016). Estimating the impact of private tutoring on academic performance: Primary students in Sri Lanka. *Education Economics, 25*(2), 142–157.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological), 39*(1), 1–22.

Dunkin, M., & Biddle, B. (1974). *The study of teaching*. New York: Holt, Rhinehart & Winston.

Fernandez, C., Yoshida, M., & Stigle, J. (1992). Learning mathematics from classroom instruction: On relating lessons to pupils' interpretations. *The Journal of the Learning Sciences, 2,* 333–365.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica, 77*(3), 953–973.

Fredricks, J. A. (2012). Extracurricular participation and academic outcomes: Testing the over–scheduling hypothesis. *Journal of Youth and Adolescence, 41*(3), 295–306.

Guill, K., Ludtke, O., & Koller, O. (2019). Assessing the instructional quality of private tutoring and its effects on student outcomes: Analyses from the German National Educational Panel Study. *British Journal of Educational Psychology*. https://doi.org/10.1111/bjep.12281.

Herppich, S., Wittwer, J., Nückles, M., & Renkl, A. (2013). Does it make a difference? Investigating the assessment accuracy of teacher tutors and student tutors. *The Journal of Experimental Education, 81*(2), 242–260.

Hu, Y., Fan, W., & Ding, W. (2015). Does shadow education aggravate inequality of educational outcomes? An empirical study on PISA2012 Shanghai data. *Peking University Education Review, 13*(3), 29–46,188.

Koenker, R., & Bassett, G., Jr. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society, 46*(1), 33–50.

Kuan, P. Y. (2011). Effects of cram schooling on mathematics performance: Evidence from junior high students in Taiwan. *Comparative Education Review, 55*(3), 342–368.

Lee, J. Y. (2013). *Private Tutoring and Its Impact on Students' Academic Achievement, Formal Schooling and Educational Inequality in Korea*. (Unpublished doctoral dissertation). Columbia University, New York, USA.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2014). *IBM SPSS for intermediate statistics: Use and interpretation*. Abingdon: Routledge.

Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics, 47*(1), 35–51.

Liang, Y. (2014). Teachers' perceptions on the potential impact of shadow education. *Renmin University of China Education Journal, 4,* 89–99.

Liu, J. (2012). Does cram schooling matter? Who goes to cram schools? Evidence from Taiwan. *International Journal of Educational Development, 32*(1), 46–52.

Meyer, E., & Van Klaveren, C. (2013). The effectiveness of extended day programs: Evidence from a randomized field experiment in the Netherlands. *Economics of Education Review, 36,* 1–11.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College. Mullis, I. V. S.

National Center of Education Quality. (2018). *National report on Chinese Compulsory Education*. http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/201807/t20180724_343663.html.

Ömeroğulları, M., Guill, K., & Köller, O. (2020). Effectiveness of private tutoring during secondary schooling in Germany: Do the duration of private tutoring and tutor qualification affect school achievement? *Learning and Instruction, 66,* 101306.

Park, H., Buchmann, C., Choi, J., & Merry, J. J. (2016). Learning beyond the school walls: Trends and implications. *Annual Review of Sociology, 42*(1), 231–252.

Smyth, E. (2009). Buying your way into college? Private tuition and the transition to higher education in Ireland. *Oxford Review of Education, 35*(1), 1–22.

Stein, K. M., Grover, W. B., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal Summer, 33*(2), 455–488.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal, 113,* 3–33.

Wang, L., & Cao, Y. (2014). Literature review of mathematics teacher effect on student achievement. *Journal of Mathematics Education, 23*(3), 48–52.

Wang, L., & Guo, K. (2017). Shadow education of mathematics in China. In Y. Cao & F. K. S. Leung (Eds.), *The 21st century mathematics education in China* (pp. 93–103). New York: Springer.

Wang, J., & Lin, E. (2005). Comparative studies on U.S. and Chinese mathematics learning and the implications for standards–based mathematics teaching reform. *Educational Researcher, 34*(5), 3–13.

Wang, L., Cao, Y., & Guo, K. (2018). A research on mathematics teachers' effect on students' achievement. *Teacher Education Research, 30*(1), 87–94.

Wang, L., Gong, X., & Pei, C. (2019). Shadow education research with mathematics education: Its value and future research area. *Journal of Mathematics Education, 28*(1), 79–82.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 390–408.

Yung, K. W. (2020). Comparing the effectiveness of cram school tutors and schoolteachers: A critical analysis of students' perceptions. *International Journal of Educational Development, 72,* 102141.

Zeng, X., & Zhou, H. (2012). A descriptive analysis on after–school tutoring in Beijing: Its costs and benefits. *Journal of Educational Studies, 8*(6), 103–109.

Zhang, Y. (2011). *The determinants of National College Entrance Exam performance in [People's Republic of] China—With an analysis of private tutoring. (Unpublished doctoral dissertation).* Columbia University, New York, USA.

Zhang, Y. (2013). Does private tutoring improve students' national college entrance exam performance? A case study from Jinan, China. *Economics of Education Review, 32,* 1–28.

Zhang, Y., & Liu, J. (2016). The effectiveness of private tutoring in China with a focus on class–size. *International Journal of Educational Development, 46,* 35–42.

Zhang, Y., Ma, X., & Wang, L. (2020). The determinants of private tutoring participation for mathematics in China: Focusing on the role of student metacognition. *Frontiers in Psychology.* https://doi.org/10.3389/fpsyg.2020.00603

Zhao, G. (2015). Can money "buy" schooling achievement? Evidence from 19 Chinese cities. *China Economic Review, 35,* 83–104.

Zheng, X., Wang, C., Shen, Z., & Fang, X. (2020). Associations of private tutoring with Chinese students' academic achievement, emotional well–being, and parent–child relationship. *Children and Youth Services Review.* https://doi.org/10.1016/j.childyouth.2020.104934