



Does private supplementary tutoring matter in Chinese students' learning of mathematics: a longitudinal study

Yuhuan Zhang¹ · Chengcheng Cui¹ · Yahan He² · Lidong Wang^{3,4}

Accepted: 28 February 2022 / Published online: 17 March 2022
© FIZ Karlsruhe 2022

Abstract

Most parents consider private supplementary tutoring (PT) as an effective means of improving academic achievement. However, previous studies on this method have produced either partial or inconclusive results on its effectiveness. Thus, in the present study we conducted a comprehensive analysis of the learning of middle school students in China, with specific focus on the final year of middle school. The analysis was based on a specially designed longitudinal survey of private supplementary tutoring in mathematics. An analysis using hierarchical linear regression showed that regular private tutoring, throughout the school years, could have a minor effect on students' mathematical achievements by the final year of middle school. These results suggest that parents should make careful choices for their children's private tutoring, and the government must issue comprehensive, professional guidelines to regulate the private tutoring industry. Moreover, other countries could find major take-aways from the Chinese experience of private tutoring for enhancing students' mathematical performances.

Keywords Private supplementary tutoring · Hierarchical linear modelling · Mathematical achievement · China · Effectiveness

1 Introduction

Private supplementary tutoring (PT) refers to the provision of paid, organised, and structured lessons on academic subjects, outside of formal school, while mimicking the formal school curriculum (Bray, 2009; Kuan, 2011; Zhang & Liu, 2016). PT has been reported as being especially popular in East Asian countries, such as China, Japan, and Singapore, partly because of the tense competition in examinations with high stakes, and partly because their culture traditionally holds education in high regard (Byun & Baker, 2015; Byun et al., 2018; Ömeroğulları et al., 2020; Wang & Guo, 2017; Wang et al., 2014). Data from the Chinese National

Assessment for Education Quality, an authorised national survey on education, documented that 43.8% of fourth-grade and 23.4% of eighth-grade students participated in mathematics shadow education in 2015, and approximately 45% engaged in PT outside school for more than two hours per week (National Centre of Education Quality, 2018). Over the last two decades, PT has gained popularity in regions outside of East Asia as well and attracted attention from researchers worldwide, including in Europe and the United States (Bray, 2014; Davies, 2004; Mischo & Haag, 2002; Zhang et al., 2021). This study highlights the experience of PT in the Chinese context, for the purpose of gaining insight into the scope of PT internationally.

Although PT might have different characteristics in different cultures, the education research community agrees on the fact that it has the potential to exert significant influences on teaching and learning in schools. Internationally, consumers' aspirations for PT to enhance students' academic achievements have made it an educational phenomenon as important as formal school education. Given that PT consumes significant family resources worldwide, stakeholders have clear questions on whether it is successful in improving students' academic achievements (Bray, 2014), and its functioning in the context of daily learning for making these improvements.

✉ Lidong Wang
wldyaoyuan@163.com

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

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

³ Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University, Beijing, China

⁴ International Center for Research in Mathematics Education, Beijing Normal University, Beijing, China

Academic achievement assessments in mathematics, conducted internationally, suggest that Chinese and other East Asian students have clear advantages over their Western peers (Leung, 2001, 2018; Mullis et al., 2004; Wang & Lin, 2005). Typically, this advantage has been explored from the perspective of classroom instruction in formal schooling (Leung, 2018), with a plethora of studies that compare Chinese and Western cultures in this regard (Clarke et al., 2006, 2012). However, this could be attributed to the possibly greater efforts of Chinese students, both in formal learning in school and PT outside school (Bray et al., 2020; Wang & Guo, 2017). Therefore, it is important to explore the effect of PT on students' mathematical achievements, when studying mathematical education in China (Wang & Guo, 2017; Zhang et al., 2021).

The effect of PT on students' achievements should be clarified for the purpose of determining whether the high rate of PT participation is an important factor in the comparatively greater academic achievements of Chinese and other East Asian students. Moreover, results from the Chinese context might have implications for other countries with fast-growing PT participation rates. With regard to the appropriate methodology for evaluating PT effectiveness, there is a marked absence of experimental work, except that of Meyer and Van Klaveren (2013), which examined the effectiveness of extended day programs through a randomised field experiment. Longitudinal research must be conducted for stronger data analysis on causal factors, rather than relying solely on correlation analysis (Bray, 2014). While cross-sectional data have been insufficient for identifying the causal impact of PT on students' academic achievements, longitudinal studies have focused on the contribution of PT to students' achievement gains, instead of their achievement at any given time.

Therefore, in the present study we evaluate the effect of PT on mathematical achievement in China, based on a specifically designed longitudinal database, and provide insights into the 'secret' behind the greater mathematical achievements and the 'Chinese experience' of PT in mathematics, for the benefit of the international educational society.

2 Background

2.1 PT in the Chinese context

Chinese education has been traditionally influenced by Confucian values, which hold it in high regard (Wang et al., 2014). With respect to career-planning, it is valued as appropriate preparation for promising jobs and fulfilling lives, which serves as external motivation for education (Leung, 2001; Wong et al., 2012). Moreover, the 'examination culture' (Wong et al., 2012) demonstrates that examinations have an important status and high stakes in China (Wang

et al., 2014). Chinese parents and students tend to attribute greater achievement to hard work, rather than talent, and study is considered a serious endeavour in which students must expend great effort and perseverance (Leung, 2001). Such a cultural setting potentially motivates Chinese parents to seek out PT as a learning resource that can aid their children in achieving more. In general, PT in mathematics is a common social phenomenon in Mainland China (Wang & Guo, 2017). Parents tend to feel pressure to invest in their children's education during the early academic years, and PT, alongside formal schooling, presents itself as a profitable avenue for such investment (Zhang, 2018).

A national level survey (National Centre of Education Quality, 2018) showed consistently high participation rates in mathematics PT, especially at the compulsory and lower grade levels, in big cities such as Beijing and Shanghai (Wang & Guo, 2017). Moreover, participation rates in mathematics PT were consistently and considerably higher than those for other school subjects. The only comparable subject was English (the second language in the Chinese school system). Mathematics has clearly emerged as the most important subject in the shadow education system; this is consistent with the findings of studies conducted in other countries (Bray & Lykins, 2012; Byun, 2014; Guill et al., 2019).

Especially in the larger cities, where PT is very prevalent, some schools find students who have already learned their lessons at an informal tutoring service, before their teaching of the content (Chen & Chen, 2015). In fact, some teachers express concern about PT possibly becoming the natural process of learning mathematics (Liang, 2014) and posing unnecessary challenges to mathematics teachers in schools.

This review concludes that PT has a vast scale, involving large numbers of students from all grade levels (He et al., 2021). These students and their parents spend long hours and considerable money on PT (especially in big cities and at compulsory levels). A sizeable portion of students' study hours is spent on PT. The high participation rate in China provides abundant information on the characteristics of PT instruction, effectiveness, etc. especially for countries that are witnessing higher PT participation rates than ever before.

2.2 Theoretical considerations concerning mathematics learning relating to PT in China

Theories connecting PT to students' mathematical learning have not advanced adequately (Guill et al., 2019; Zhang et al., 2021). Carroll's model of school learning (1963), which posits learning as a function of the ratio of the time spent on learning to the time required to learn, provides a theoretical foundation for understanding the role of PT in students' mathematical learning (Guill et al., 2019; He et al.,

2021; Zhang et al., 2021). By adding extra learning hours, PT potentially enhances students' understanding; thus, time must be taken into consideration when examining the factors that impact mathematical learning and achievements. Previous studies have not addressed PT suitably in the analysis of factors impacting academic achievement (see the review by Winne & Nesbit, 2010).

In addition, Carroll's model (1963) states that simply increasing learning time might not sufficiently enhance academic achievement, as the quality of PT instruction is also important for effectiveness. Therefore, the quality of PT should be considered when exploring students' mathematical learning (achievement), especially when the participation rate is quite high, such as in Eastern countries. It is not enough to simply assume that PT plays an important role; it requires investigation. In the current study we tested the hypothesis that both formal-school and out-of-school mathematics instruction can impact students' mathematical learning. The following conceptual framework (He et al., 2021) served as a guide for the purpose.

The two modes of instruction, i.e., formal-school and out-of-school, potentially interact with each other. Both could be impacted by student-level variables, such as the socio-economic status of the family, a student's interest in mathematical learning, and availability of computers and books in the family (Winne & Nesbit, 2010; Zhang et al., 2020) (Fig. 1).

For the covariates controlling the endogenous bias in the PT effect, besides the use of several predictors (e.g., gender, and urban areas; Byun et al., 2018; Smyth, 2009) previous studies tended to use 'popular' variables in large-scale surveys, such as family SES (e.g., Kuan, 2011) or parents' education and educational expectation (e.g., Liu, 2012). Students' basic level of achievement was found to be a core factor in their PT participation (Baker et al., 2001; Zeng & Zhou, 2012), while parents' education,

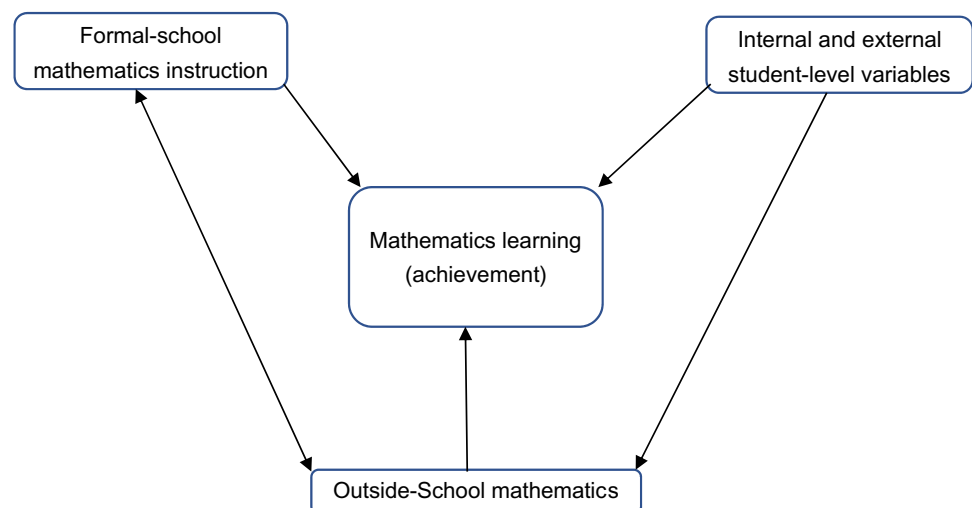
pre-intervention score in mathematics, and educational expectations (e.g. Kuan, 2011; Liu, 2012; Stevenson & Baker, 1992) were identified as control variables. In addition, classroom factors were identified as having a fundamental impact on students' achievements (Winne & Nesbit, 2010). These were not comprehensively included in previous studies on PT effect. Thus, a conceptual framework comprising variables at the levels of individual, family and class level should guide the analysis of PT effectiveness.

2.3 Measuring the effect of PT on mathematical achievement in China

As part of research on educational effectiveness, quantitative studies have evaluated the effect of PT on students' performance in mathematics. Typically, researchers have used a regression model, by controlling the selection bias of PT and endogeneity (Zhang, 2013), for achieving a partially causal model. While cross-sectional data have been inefficient in identifying the causal impact of PT on students' academic achievements, longitudinal studies have mostly focused on the contribution of PT to students' achievement gains (Zhang et al., 2021). Therefore, detailed data analysis of the causal factors is still required in longitudinal research (Bray, 2014).

Typically, in the measurements of PT, whether or not students received PT is used as a dummy variable (Zhang, 2013) for comparing the achievements of tutored and non-tutored students (e.g., Baker et al., 2001; Bray, 2014; Byun, 2014; Guill et al., 2019; He et al., 2021; Kuan, 2011; Ömeroğulları et al., 2020; Zhang, 2013; Zhang et al., 2021). More recent studies have explored the effect of this dummy variable on mathematical achievement in China, based on different kinds of databases (e.g., Fang et al., 2018; Guo et al., 2020; Sun et al., 2020; Zeng & Zhou, 2012; Zhang,

Fig. 1 Conceptual framework



2011, 2013; Zhang et al., 2021; Zhao & Xue, 2021; Zheng et al., 2020). However, only three of these were systematic, retrospective longitudinal studies. Zhang (2011, 2013) did not find a significant correlation between attending mathematics PT sometimes during the final year of high school, and mathematical performance in the Chinese National College Entrance Exam, even after controlling for the total score of the entrance exam and other covariates. Zhang et al. (2021) also did not find any significant positive effect of regular PT attendance on mathematical achievements at the middle school level, even with continuous PT participation throughout the academic year (not the final year). While He et al. (2021) did find a significant positive effect of regular PT attendance during summer vacation on mathematical achievements at the high school level, they did not find a significant positive effect of attending PT during the school semester.

It must be noted that simply inquiring whether a student received PT is too broad a question for meaningful statistical analyses (Bray, 2014). Factors such as the point in time at which PT was received, the amount of time spent on it (e.g., one semester or one academic year), and systematic or random attendance (He et al., 2021; Ömeroğulları et al., 2020; Zhang et al., 2021) must also be considered, since they might have a substantial impact on the statistical results of the evaluation of PT effectiveness. Thus, ‘Does private supplementary tutoring work?’ is too broad to be a meaningful question (Bray, 2014), and related studies require a more careful research design, especially including the detailed measurement of PT.

2.4 Research questions

In the present study we focused on the following research questions:

1. What are the PT participation rates across different periods of (almost) a year in China?
2. Does regular PT participation for (almost) a year have an effect on students’ mathematical achievements?

3 Method

3.1 Participants

Data were collected in a medium-sized city in central China, with intermediate levels of economic and educational development. Two rounds of longitudinal data were collected through a questionnaire survey, conducted in December 2019 and January 2021. Administrative data were obtained from the head of the department of mathematics at the city’s education bureau. The survey tracked PT throughout

the second semester of the eighth grade, summer vacation, and the first semester of the ninth grade. Additionally, final mathematics scores for the semesters were collected from the schools. To ensure the validity and accuracy of the survey, the city’s middle schools were divided into three levels, under the supervision of the head of the department of mathematics, because they differed in their quality of schooling. A total of five schools (the number of first, second, and third level schools was 2, 2, and 1, respectively), 52 classes and 2645 students were selected using whole-level sampling. The students answered the questionnaire on their own, under the supervision of the class teacher, and the end-of-semester final examinations were organised by the city’s education bureau. After matching the two data sets with the names of students in the class, data with more than 20% of missing values were excluded, resulting in 2123 valid responses. The proportions of students from the first, second, and third level schools in the final sample were 32.6%, 36.9% and 30.4%, respectively.

3.2 Missing data

The imputation method of Expectation Maximization (EM) in SPSS 23.0, was used to impute the missing data. It is a number algorithm that can be used to maximise likelihood under a wide variety of missing-data models (Dempster et al., 1977).

3.3 Measures

3.3.1 PT participation

The questionnaire questioned students primarily on their mathematics PT participation at the following levels: the second semester of eighth grade, the summer vacation and the first semester of ninth grade. They were asked about their PT intensity (never, occasionally, or on an average of 1 h, 2 h, 3 h, 4 h or more each week). In this study, we coded only two situations: 1 = participation, 0 = no participation (including ‘occasionally’). Combinations of PT in the different periods were coded as follows: 000 (no PT in any of the three periods), 100 (PT only during eighth grade), 010 (PT only during summer vacation), 001 (PT only during ninth grade), 110 (PT during summer vacation and eighth grade), 101 (PT in eighth and ninth grades), 011 (PT during summer vacation and ninth grade) and 111 (PT in all three periods).

3.3.2 Outcome variables

The dependent variable was the mathematics score of students at the end of the first school semester of ninth grade.

The test was designed by the head of the department of mathematics, to assess students' mathematical learning in the first school semester of ninth grade; thus, errors caused by inconsistent measuring tools were avoided (Zhang et al., 2021). The total raw score was 120, which was standardised into z-scores across five schools, to eliminate the dimensional relationship between variables and make the data more comparable.

3.3.3 Covariates

Achievement at school is influenced by individual, familial, and classroom factors. Individual factors include the student's gender, self-expectation of education, pre-score at mathematics, i.e., the final mathematics examination result of the first semester of eighth grade (e.g., Byun et al., 2018; Liu, 2012). Familial level factors include mother's education, parents' highest occupation (Li, 2003), availability of a personal learning desk at home (e.g. Byun et al., 2018; Kuan, 2011) etc. The class level factors include whether the student is in a key class, the head teacher's management

style, classroom learning atmosphere, and whether the head teacher is a mathematics teacher. The controlled covariates are described in Table 1.

3.4 Data analysis

Students in the same class might be affected by different characteristics. In this study we used a two-level Hierarchical Linear Model (HLM) for evaluating the effect of PT participation on students' mathematical performance. The first level is that of the student and the second level is that of the class, and the students are nested below the class. The empty model does not contain any student or class variables; it is used to measure differences in students' mathematics scores across classes (if any).

$$y_{ij} = \beta_{0j} + \beta_{1j}PT_{ij} + \beta_{2j}X_{ij} + \varepsilon_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{0s}K_j + \mu_{0j} \quad (2)$$

Table 1 Definitions and measures of the variables

Variable	Definitions and measures
<i>Dependent variable</i>	
Mathematics achievement	Mathematics achievement scores obtained at the end of the first school semester of ninth grade (z-score, total full score of raw score: 120)
<i>Independent variables</i>	
Participation	Control group: 000 (no PT in any of the three periods) Treatment groups: 100 (PT during eighth grade only), 010 (PT during summer vacation only), 001 (PT during ninth grade only), 110 (PT during summer vacation and eighth grade), 101 (PT during eighth and ninth grades), 011 (PT during summer vacation and ninth grade), 111 (PT in all three periods), where 1 = participation and 0 = no participation
<i>Covariates</i>	
<i>Individual level</i>	
Gender	Male = 1, female = 0
Only child	Yes = 1, no = 0
Self-educational expectation	senior high school or lower = 1; junior college = 2; undergraduate = 3; master = 4; PhD = 5
Difficulty in learning	Particularly difficult = 1, a little difficult = 2, generally difficult = 3, not very difficult = 4, not difficult at all = 5
Pre-score in mathematics	Mathematics achievement scores received at the end of the first school semester of eighth grade (z-score, total full score of raw score: 100)
<i>Family level</i>	
Mother's education	Primary school or less = 1; junior high school = 2; technical secondary school, vocational high school, senior high school = 3; junior college, undergraduate = 4; Masters, PhD = 5
Parents' occupation	Three categories: lower level (worker, peasant, primary worker, individual businesses, retirement) (reference group); middle level (accountant, nurse, engineer, clerk, service personal, skilled worker); upper level (leader of government agency or company, professionals, doctor, lawyer, teacher)
Parents' educational expectation	senior high school or lower = 1; junior college = 2; undergraduate = 3; master = 4; PhD = 5
Books owned by family	Very few = 1, few = 2, moderate = 3, multiple = 4, many = 5
<i>Class level</i>	
Whether it is key class	Yes = 1, no = 0
Classroom learning atmosphere	Least promotive = 1, less promotive = 2, average = 3, more promotive = 4, most promotive = 5

Here, y_{ij} is the grade of i from class j , β_{0j} is the average grade of the j class, β_{1j} is the regression coefficient of the independent variable, β_{2j} is the regression coefficient of the covariate, ε_{ij} is the unexplained residual of the score of i from class j , γ_{00} is the average grade of all students, γ_{0s} is the regression coefficient of class-level variable, K_j is the class level variable, and μ_{0i} is the unexplained residue at the class level. Before running the HLM analysis, we examined the variance under the empty models, and found that the variance at the student level was 0.667, and that between classes was 0.336. The intra-class correlation coefficients (ICC) under the empty models for students' standardised test scores showed strong nesting effects, with an ICC of 0.335. The results showed that multilevel modelling is necessary for this study.

4 Results

4.1 Descriptive statistics and correlation analysis

Table 2 displays the descriptive statistics for all study variables. The proportions of students participating in mathematics PT in eighth grade, summer vacation and ninth grade were 38.8%, 48.1% and 49.5%, respectively. The proportion of students who did not participate in PT in any of the three periods was 39.5%, that of students who participated in PT in only one of the three periods was 1.7% (eighth grade), 5.9% (summer vacation), and 7.7% (ninth grade), and that of students who participated in PT in all three periods was 30.6%.

Table 2 Descriptive statistics of the variables used for analysis

Name of variable	<i>N</i>	%	Name of variable	<i>N</i>	%
<i>Independent variables</i>					
<i>Participation</i>			Lower level	608	28.6
PT in eighth grade	823	38.8	Middle level	750	35.3
PT in summer vacation	1021	48.1	Upper level	765	36.0
PT in ninth grade	1050	49.5	<i>Mother's education</i>		
000	838	39.5	Primary school or less	134	6.3
100	36	1.7	Junior high school	665	31.3
010	126	5.9	Technical secondary school, vocational high school, senior school	817	38.5
001	163	7.7	Junior college, undergraduate	458	21.6
110	73	3.4	Master's, PhD	49	2.3
101	65	3.1	<i>Books owned by family</i>		
011	173	8.1	Rarely	76	3.6
111	649	30.6	Less	144	6.8
<i>Covariates</i>			Average	791	37.3
<i>Individual level</i>			More	744	35.0
Male	1076	50.7	A lot	368	17.3
Only child	480	22.6	<i>Parents' educational expectation</i>		
<i>Difficulty in learning mathematics</i>			Senior high school or lower	181	8.5
Particularly difficult	254	12.0	Junior college	156	7.3
A little difficult	689	32.5	Undergraduate	843	39.7
Generally difficult	683	32.2	Master	362	17.1
Not very difficult	370	17.4	PhD	581	27.4
Not difficult at all	127	6.0	<i>Class level</i>		
<i>Self-educational expectation</i>			Key class	840	39.6
Senior high school or lower	280	13.2	<i>Class learning atmosphere</i>		
Junior college	163	7.7	Least promotive	71	3.3
Undergraduate	797	37.5	Less promotive	153	7.2
Master	425	20.0	Promotive	666	31.4
PhD	458	21.6	More promotive	724	34.1
<i>Family level</i>			Most promotive	509	24.0
<i>Parents' occupation</i>			Total	2123	100

Fig. 2 Dendrogram of the proportion of students participating in PT in the three periods

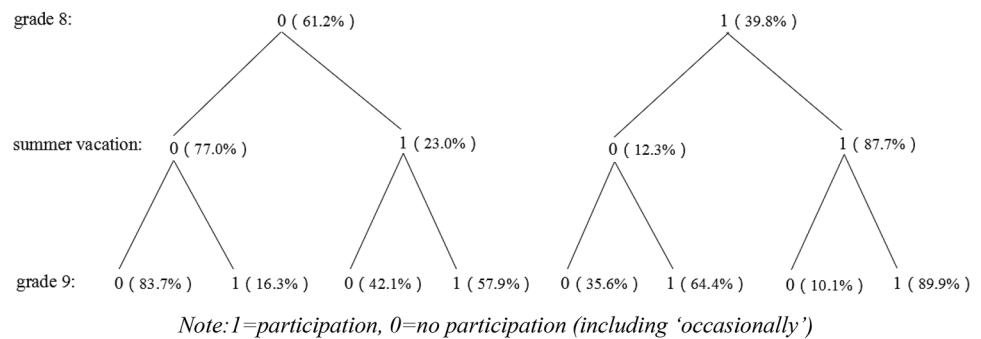


Figure 2 shows participation in PT in the three periods. One can observe that students show a certain degree of continuity in PT participation.

4.2 HLM analysis

We established Model 1 as a two-level model without any variables, except for pre-score in mathematics. It shows the extent to which the variance (46%) is explained by the variable of pre-score at mathematics, thereby demonstrating the benefits of the longitudinal design. To prevent overfitting, we used the manual stepwise regression for picking the most relevant variables for establishing Model 2. To make the model more realistic, adding relevant variables based on suggestions from previous studies on the basis of model 2 is model 3. The results of the HLM analysis indicated that, after controlling for the students' individual, family and class level variables, regular PT attendance throughout the three periods had a significant effect on students' mathematical achievement at the end of the first school semester of the final year of middle school, with a small effect size (0.108, $p = 0.020$). In addition, PT in ninth grade played an important role, as all the combinations that included PT participation in ninth grade seemed to positively impact students' mathematical achievements, with statistical or marginal significance. However, 101 and 011 did not have any significant effect. Moreover, students who participated in PT during summer vacation, and quit it in the ninth grade, seemed to have a lower mathematical achievement at the end of the first school semester of ninth grade, with marginal significance.

Of all the covariables, self-expectation of education, parents' expectation of their children's education, and being in a key class had significant impacts.

For Models 2 and 3, the variance explaining the rate of level 1 variables is $f^2 = \frac{0.667-0.304}{0.667} = 0.543$, and the variance explaining rate of level 2 variables is $f^2 = \frac{0.336-0.092}{0.336} = 0.725$. These results show that the variables have a strong effect (Cohen, 1992) (Table 3).

5 Discussion

In this study we conducted a longitudinal analysis of the effectiveness of long-term regular PT attendance on students' mathematical achievements in the final grade of Chinese middle school.

5.1 Scale of PT participation in China

The results of this study support those of previous studies, showing that approximately 40% of students regularly participated in mathematics PT, even during the COVID-19 pandemic. In addition, the longitudinal database reflected the variations in PT participation across three periods in one year. Students tend to participate continuously or not participate in PT during different periods in the same year (PT in the second half of the academic year of eighth grade, summer vacation, and the first half of ninth grade); thus, PT participation can be considered a 'habit' among some students, and PT participation in a particular period influences participation in the successive periods (Zhang et al., 2020). Future studies could focus on how families decide on PT participation.

5.2 The effect of PT on mathematical achievement in China

The results of the HLM analysis indicate that regular PT attendance throughout a year could enhance students' mathematical achievements by the middle of the final year of middle school, with a small effect size.

According to Carroll's (1963) model of school learning, extending the time spent on PT (up to a whole academic year) could enhance PT effectiveness; however, PT participation for half a year was not found to be effective (Zhang et al., 2021). It must be noted that the COVID-19 pandemic struck China during the spring of 2020, and in the city where data were collected, students returned to school after two months, i.e., in May. Hence, the effect of PT might have been

Table 3 Result from the HLM analysis for PT effectiveness

Variable	Model 1	Model 2	Model 3
<i>Independent variables</i>			
Private tutoring			
000 (control group)			
100		0.139(0.087)	0.131(0.088)
010		- 0.009(0.053)	- 0.010(0.053)
001		0.144**(0.053)	0.147**(0.054)
110		- 0.010(0.052)	- 0.005(0.052)
101		0.162(0.092)	0.168(0.091)
011		0.075(0.057)	0.078(0.057)
111		0.102*(0.046)	0.108*(0.047)
<i>Covariates</i>			
Individual level			
Gender			- 0.006(0.029)
Only child			0.036(0.037)
Difficulty in learning mathematics		0.185***(0.028)	0.187***(0.028)
Self-educational expectation		0.063***(0.014)	0.064***(0.013)
Pre-score in mathematics	0.731***(0.041)	0.552***(0.053)	0.551***(0.053)
Family level			
Mother's education			- 0.007(0.015)
Parents' occupation			
Lower level (reference group)			
Middle level			- 0.029(0.032)
Upper level			- 0.034(0.038)
Books owned by family			- 0.020(0.014)
Parents' educational expectation		0.052***(0.013)	0.054***(0.013)
Class level			
Key class		0.182**(0.064)	0.182**(0.182)
Class learning atmosphere		0.104*(0.043)	0.105*(0.042)
Intercept	- 0.091***(0.049)	- 1.408***(0.236)	- 1.326***(0.226)
Random effect			
Variance component			
Level-1 variance	0.360	0.304	0.304
Level-2 variance	0.050	0.020	0.020
Chi-square	337.87***	177.45***	176.32***
Level-1 f^2		0.543	0.543
Level-2 f^2		0.725	0.725
Empty model ICC coefficient	0.335		

* < 0.05, ** < 0.01, *** < 0.001. Standard errors are shown in brackets

enhanced by the suspension of formal school instruction during the pandemic. Therefore, it is likely that the small PT effect might be obliterated after the pandemic. However, these results compensate for the limitations of previous studies, which typically used PT as a dummy variable and indicated that PT might be effective alongside formal schooling, such as in scenarios like a pandemic.

In addition, several other variables, such as PT participation during the final year of middle school, could have enhanced academic performance by the final year, when

students took the high-stake examination for entrance to well-ranked high schools. This indicates that students in the final academic year tend to be more motivated and make better use of PT, thus functioning as a moderating variable between PT and mathematical achievement. This hypothesis requires further testing.

This study shows how the evaluation of PT effectiveness is a rather complex issue (Guill et al., 2019; Zhang et al., 2021), as the measurement of PT participation could vary on several aspects, for example, the point of time at which

PT was received, the amount of time spent on it, and whether students are in their final year of middle school. Moreover, this study derived its results from a medium-sized city in central China, with a moderate level of economic and educational development. Therefore, replicating it for larger or smaller cities, with higher or lower levels of development in formal and informal schooling, requires the exercise of caution.

5.3 PT and mathematics learning among Chinese students

As mentioned in Sect. 1, the evaluation of PT effectiveness should be discussed in the context of the scope of mathematical learning, especially in East Asian countries like China with high PT participation rates.

This study partly supports the statement made in the introductory section, that PT should be considered in the exploration of mathematical learning (i.e., greater mathematical achievement) in the Chinese context. However, while previous studies did not find an overall positive effect of PT on students' mathematical achievements (Zhang, 2011, 2013; Zhang et al., 2021), this study found that regular PT attendance throughout the year had only a small effect on students' mathematical achievements in the final year of middle school. Therefore, by considering the results of previous studies as well, one could state that the high PT participation rate is not a key factor in the greater mathematical achievements of Chinese students, internationally. Rather, it could be attributed to formal school instruction.

5.4 Strengths and limitations

A major strength of this study is the analysis of longitudinal data sets specific to out-of-school mathematical learning, involving a large sample size of hundreds of tutored students in the final year of middle school. Longitudinal data allowed controlling for systematic differences between tutored and non-tutored students, and those among related control variables. Moreover, the information available on PT was quite precise, and included the intensity of PT (Zhang et al., 2021) in three consecutive periods, the lack of which was a limitation of previous studies (Ömeroğulları et al., 2020). Achievement in mathematics was generally consistent with school curriculum and PT instruction. This allowed controlling for students' actual level of knowledge exactly before and after tutoring, which solved the problem of methodology found in previous studies (Bray, 2014).

The limitation of this study is the sample size of five schools in a city, which partly limited the applicability of the results. Furthermore, it focused on PT in different periods,

rather than on strengthening PT, which would require the use of advanced statistical methods. Finally, the study neither discussed the quality of PT, such as the qualification of the tutors (Wang & Wang, 2021), nor other variables that reflect the quality of PT instruction (Zhang et al., 2021). Future studies could survey students from different areas for a detailed analysis of the effects of different types of PT. They could also pay more attention to the quality of PT, for a more comprehensive understanding of PT effectiveness.

Moreover, it cannot be definitely stated whether the small PT effect found during the COVID-19 pandemic would continue to exist after it ends, since no control group in the study was unaffected by the pandemic. However, since the sample city did not suffer severely due to the pandemic, students returned to school comparatively early. We could thus argue that the small PT effect might outlast the pandemic, but further investigation is required to support this statement.

6 Conclusion and implications

In this study we analysed the results of a specially designed longitudinal survey of private supplementary mathematics tutoring among middle school students in China. We concluded that regularly attending PT throughout could have a small effect on students' mathematical achievements by the final year of middle school. Moreover, the effectiveness of PT in the final academic year depends on students' motivation for learning. The study indicates that parents should carefully select PT for their children since it might not always yield the expected results. Mathematics PT does provide students with additional opportunities for enhancing knowledge, but this does not imply that participation in mathematics PT will necessarily have a positive impact on students' achievement. PT participation should be carefully considered by parents, in consultation with schoolteachers, and based on students' personal characteristics such as performance foundation, knowledge lacunae, grade levels, and personal PT needs. The opinions of individual students must be respected while acknowledging the unique advantages of tutoring for personalised education. In addition, the study suggests that the government must issue comprehensive and professional guidelines for regulating the large, but mostly unregulated, PT industry in China. Such guidelines would improve the standards of implementing local PT policies, clarify the responsible stakeholders, recognise the vitality of local regulatory agencies in the policy implementation process, and positively direct the development of PT institutions. Finally, the study's findings provide important insight to Western countries with significantly expanding PT rates, on the prudent use of PT in improving students' mathematical achievements.

Funding This study was Funded by Beijing Philosophy and Social Sciences (Grant no. 18JYC025).

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. <https://doi.org/10.3102/01623737023001001>
- Bray, M. (2009). *Confronting the shadow education system: What government policies for what private tutoring?* 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. <https://doi.org/10.1007/s12564-014-9326-9>
- 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*. 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 Republic of Korea. In H. Park & K. Kim (Eds.), *Korean education in changing economic and demographic contexts* (pp. 39–58). Springer. https://doi.org/10.1007/978-981-4451-27-7_3
- Byun, S. Y., & Baker, D. (2015). Shadow education. In R. Scott, S. Kosslyn, & N. Pinkerton (Eds.), *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource* (pp. 1–9). Wiley-Blackwell. <https://doi.org/10.1002/9781118900772.etrds0301>
- Byun, S. Y., Chung, H. J., & Baker, D. P. (2018). Global patterns of the use of shadow education: Student, family, and national influences. *Research in the Sociology of Education*, 20, 71–105. <https://doi.org/10.1108/s1479-35392018000020004>
- Carroll, J. B. (1963). A model of school learning. *Teacher College Record*, 64, 723–733.
- Chen, C., & Chen, Y. (2015). 中小學生課外學習的應試性特點與問題 [The characteristics and problems of extracurricular study of primary and middle school students] *Educational Research and Experiment*, 5, 43–48.
- Clarke, D. J., Emanuelsson, J., Jablonka, E., & Mok, I. A. C. (Eds.). (2006M). *Making connections: Comparing mathematics classrooms around the world*. Sense.
- Clarke, D. J., Wang, L., Xu, L., Aizikovitch- Udi, E., & Cao, Y. (2012). International comparisons of mathematics classrooms and curricula: The validity-comparability compromise. In T. Y. Tso (Ed.), *Proceedings of the 36th conference of the international group for the psychology of mathematics education (PME36)*, Taipei, Taiwan (vol. 2, pp. 171–178).
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Davies, S. (2004). School choice by default? Understanding the demand for private tutoring in Canada. *American Journal of Education*, 110(3), 233–255. <https://doi.org/10.1086/383073>
- 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. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>
- Fang, C., Hu, Y., & Zhang, P. (2018). 小学生能从课后学习时间中受益吗 [Can pupils of primary schools benefit from the study time after class?]. *Journal of Educational Science of Hunan Normal University*, 17(1), 69–77.
- 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*, 90(2), 282–300. <https://doi.org/10.1111/bjep.12281>
- Guo, Y., Chen, Q., Zhai, S., & Pei, C. (2020). Does private tutoring improve student learning in China? Evidence from the China Education Panel Survey. *Asia and the Pacific Policy Studies*, 7(3), 322–343. <https://doi.org/10.1002/app5.310>
- He, Y., Zhang, Y., Ma, X., & Wang, L. (2021). Does private supplementary tutoring matter? The effect of private supplementary tutoring on mathematics achievement: A longitudinal analysis. *International Journal of Educational Development*, 84, 102402. <https://doi.org/10.1016/j.ijedudev.2021.102402>
- 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. <https://doi.org/10.1086/659142>
- Leung, F. K. S. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics*, 47(1), 35–51. <https://doi.org/10.1023/a:1017936429620>
- Leung, F. K. S. (2018). Mathematics education of Chinese communities from the perspective of international studies of mathematics achievement. In Y. Cao & F. K. S. Leung (Eds.), *The 21st century mathematics education in China* (pp. 3–26). Springer. https://doi.org/10.1007/978-3-662-55781-5_1
- Liang, Y. (2014). 从教师角度看影子教育对师生之影响 [Teachers' perceptions on the potential impact of shadow education]. *Renmin University of China Education Journal* 4, 89–99.
- Li, C. (2003). 社会政治变迁与教育机会不平等——家庭背景及制度因素对教育获得的影响 [Socio-Political Changes and Inequality of Educational Opportunity] *Social Sciences in China*, 3, 86–98.
- 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.
- Mischo, C., & Haag, M. L. (2002). Expansion and effectiveness of private tutoring. *European Journal of Psychology of Education*, 17(3), 263–273. <https://doi.org/10.1007/bf03173536>
- 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*. TIMSS and PIRLS International Study Centre, Boston College.
- National Centre of Education Quality (2018). 中国义务教育质量监测报告 [National Report on Chinese Compulsory Education] http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/201807/t20180724_343663.html. Accessed 15 Mar 2022
- Ömerogullari, 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. <https://doi.org/10.1016/j.learninstruc.2020.101306>
- 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.

- Stevenson, D. L., & Baker, D. P. (1992). Shadow education and allocation in formal schooling: Transition to university in Japan. *American Journal of Sociology*, *97*(6), 1639–1657.
- Sun, L., Shafiq, M. N., McClure, M., & Guo, S. (2020). Are there educational and psychological benefits from private supplementary tutoring in Mainland China? Evidence from the China Education Panel Survey, 2013–15. *International Journal of Educational Development*, *72*, 102144. <https://doi.org/10.1016/j.ijedudev.2019.102144>
- Wang, C., & Wang, L. (2021). The mathematics teacher in shadow education: A new area of focus in teacher education. *Journal of Mathematics Teacher Education*. <https://doi.org/10.1007/s10857-021-09522-3>
- 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. <https://doi.org/10.3102/0013189x034005003>
- 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). Springer. https://doi.org/10.1007/978-3-662-55781-5_5
- Wang, L., Li, X., & Li, N. (2014). Socio-economic status and mathematics achievement in china: A review. *ZDM - the International Journal on Mathematics Education*, *46*(7), 1051–1060. <https://doi.org/10.1007/s11858-014-0617-8>
- Winne, P. H., & Nesbit, J. C. (2010). The psychology of academic achievement. *Annual Review of Psychology*, *61*(1), 653–678. <https://doi.org/10.1146/annurev.psych.093008.100348>
- Wong, N. Y., Wong, W. Y., & Wong, E. W. (2012). What do the Chinese value in (mathematics) education? *ZDM - the International Journal on Mathematics Education*, *44*(1), 9–19. <https://doi.org/10.1007/s11858-012-0383-4>
- Zeng, X., & Zhou, H. (2012). 北京市四、八年级学生课后补习的代价与收益 [A descriptive analysis on after-school tutoring in Beijing: Its costs and benefits]. *Journal of Educational Studies*, *8*(6), 103–109.
- Zhang, W. (2014). The demand for shadow education in China: Mainstream teachers and power relations. *Asia Pacific Journal of Education*, *34*(4), 436–454. <https://doi.org/10.1080/02188791.2014.960798>
- 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. <https://doi.org/10.1016/j.econedurev.2012.09.008>
- Zhang, Y. (2018). Private tutoring, students' cognitive ability and school engagement, and the formal schooling context: Evidence from middle school students in China. *Chinese Journal of Sociology*, *4*(2), 277–298. <https://doi.org/10.1177/2057150x18764859>
- Zhang, Y., Dang, Y., He, Y., Ma, X., & Wang, L. (2021). Is private supplementary tutoring effective? A longitudinally detailed analysis of private tutoring quality in China. *Asia Pacific Education Review*, *22*(2), 239–259. <https://doi.org/10.1007/s12564-021-09671-3>
- 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. <https://doi.org/10.1016/j.ijedu dev.2015.11.006>
- 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*, *11*, 603. <https://doi.org/10.3389/fpsyg.2020.00603>
- Zhao, Y., & Xue, H. (2021). 参加课外补习有助于中考升学吗——基于CFPS2010—2018追踪数据的实证分析 [Does extracurricular tutoring contribute to the entrance examination of junior high school to senior high school]. *Journal of Educational Science of Hunan Normal University*, *20*(1), 24–35.
- 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*, *112*, 104934. <https://doi.org/10.1016/j.childyouth.2020.1049>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.