The Eurasia Proceedings of
Educational & Social Sciences (EPESS)
ISSN: 2587-1730

The Eurasia Proceedings of Educational & Social Sciences (EPESS), 2016

Volume 4, Pages 11-32

ICEMST 2016: International Conference on Education in Mathematics, Science & Technology

DOES SHADOW EDUCATION AGGRAVATE INEQUALITY OF EDUCATIONAL OUTCOMES

Yongmei HU
E-Mail:icemstoffice@gmail.com
Faculty of Education, Beijing Normal University/
National Innovation Center for Monitoring the Basic Education Quality

Wenfeng FAN
Faculty of Education, Beijing Normal University

Weili DING
School of Policy Studies and department of economics, Queen's University/
Faculty of Education, Beijing Normal University

ABSTRACT: The "shadow education" system of private supplementary tutoring has become quite common in East Asian countries nowadays. Based on the data of Programme for International Student Assessment 2012 (PISA 2012), the paper analyzes the influence of shadow education on the mathematical literacy of students of Shanghai, Hong Kong, Japan and Korea by means of a hierarchical linear model, and estimates the net effect of shadow education using the method of reweighting on propensity score matching (RPSM). The following findings are obtained from the study: First, supplementary math tutoring has a significant positive effect on the math score of students, and such an effect is more obvious on Japanese and Korean students than on Hong Kong and Shanghai students; second, supplementary math tutoring and supplementary science tutoring complement each other; and third, attending supplementary math tutoring may narrow the gap between students in learning performance that is caused by the difference in their families' economic, social and cultural status (ESCS), thus promoting the equality of educational outcomes. Therefore, governments and schools are advised to provide necessary opportunities of supplementary tutoring for low-capacity students from low-income families and waive their tuition fees; large-sized extracurricular education groups should be encouraged to establish after-school learning funds and/or incentive funds for students from poor families and grant fee remissions to those from ultra-low income families, so as to create a situation where the government, the school and the society jointly promote the equality of educational outcomes in the stage of compulsory education.

Keywords: Shadow Education, Inequality of educational outcomes, PISA, RPSM

INTRODUCTION

Over the past two decades, "shadow education" (or namely private supplementary tutoring) has developed to a considerable scale in some countries, where it has become a supplement to or even a duplication of mainstream education[i]. Such private supplementary tutoring refers to chargeable teaching activities outside the formal school education system that are conducted corresponding to school courses[ii]. The forms of receiving private supplementary tutoring include hiring private tutors, or going to cram schools, etc (Xue Haiping and Ding Xiaohao, 2009). As Baker and LeTendre pointed out in 2005, shadow education is not meant to replace, but rather to supplement or support formal education. By means of supplementary tutoring after school, students want to better absorb and master the knowledge taught in class so that they can be more competitive in their classes or grades and parents expect their children to improve their examination results at school.

In various parts of the world, due to the impacts of globalization, social development indicator ranking and international student achievement assessment, education competition between different countries is being increasingly intensified, and the shadow education system of private supplementary tutoring is expanding continuously[iii] (Bray 2005). This phenomenon has been very common in East Asia, especially in China, Japan

- Selection and peer-review under responsibility of the Organizing Committee of the conference

*Corresponding author: Yongmei HU- icemstoffice@gmail.com

and Korea (Stevenson & Baker 1992). In the mid-1990s shadow education began to rise in Mainland China, and now it has become quite common in Chinese urban areas. The prevalence of shadow education in these countries has much to do with their competitive entrance examination and the deep influence of Confucianism[iv].

According to a telephone survey conducted in 2009 in Hong Kong for 521 students, 72.5% of elementary school pupils of higher grades received private supplementary tutoring (Ngai, Angela and Sharon Cheung 2010). Another survey (Caritas, Community and Higher Education. 2010) conducted for 898 Hong Kong middle school students revealed that private supplementary tutoring was received by 72.5% of lower-grade students (Grade 1 to Grade 3), 81.9% of middle-grade students (Grade 4) and 85.5% of higher-grade students (Grade 5 to Grade 7). According to a survey performed in 2007 in Japan, 15.9% of Grade-1 elementary school children attended tutorial school after class and this percentage steadily rose with the increase in school grade, reaching 65.2% for junior high school students of Grade 3 (Japan, Ministry of Education and Training. 2008). It is estimated that 87.9% of elementary school students, 72.5% of junior high school students and 60.5% of senior high school students in Korea received private supplementary tutoring in 2008 (Kim, Kyung-Keun 2010). The case of China was similar to that of Korea, with the percentage falling with the increase in school grade. Based on the 2004 education and employment statistics of Chinese urban residents, Xue Haiping, et al found that the percentage of students receiving private supplementary tutoring was respectively 73.8%, 65.6% and 53.5% for elementary, junior high and senior high school students, declining with the increase in school grade (Xue Haiping and Ding Xiaohao 2009). Based on the statistics of PISA 2009 for Shanghai, Hong Kong, Japanese and Korean students (See Table 1), we found that of all the four types of private supplementary tutoring (i.e., language tutoring, math tutoring, science tutoring, and other tutoring), math tutoring was attended by the highest percentage of students (except for Hong Kong): 71.3%, 48.5%, 76.6% and 77.5% respectively, while science tutoring was attended by the lowest percentage of students: 29%, 27.9%, 61.6% and 57% respectively. Japanese and Korean students were similar to each other in terms of the percentage for language tutoring, math tutoring, science tutoring, and other tutoring. Shanghai and Hong Kong students were much alike in terms of the percentage for science tutoring but quite different in the percentage for language tutoring and math tutoring. The percentage of Shanghai students receiving language tutoring and math tutoring was about 23 percentage points higher than that of Hong Kong students. According to the percentages for math tutoring shown in Table 1, the percentages of students receiving private supplementary tutoring in Shanghai, Hong Kong, Japan and Korea were not significantly different from the results of the above-mentioned five surveys.

Table 1: Statistics of PISA 2009 on private supplementary tutoring for Shanghai, Hong Kong, Japanese and Korean Students

Country/Region	Percentage for language tutoring [a]	Percentage for math tutoring	Percentage for science tutoring [b]	Percentage for other tutoring [c]
Shanghai, China	54.3%	71.3%	29.0%	62.1%
Hong Kong, China	30.6%	48.5%	27.9%	55.3%
Japan	64.4%	76.6%	61.6%	76.0%
Korea	68.0%	77.5%	57.0%	67.8%

Notes: a. Language tutoring refers to native language tutoring. For example, language tutoring for Shanghai students refers to Chinese tutoring. b. Science tutoring refers to tutoring in science-related courses of junior high school. It covers such courses as physics, chemistry and biology for Shanghai students, for example. c. Other tutoring refers to tutoring in courses not listed in the first three columns. For Shanghai students as an example, English tutoring is classified as other tutoring.

In China and some other countries, the expenditure on private supplementary tutoring has become an important part of a family's education spending (Li & Tsang 2003; Tsang 2002). According to a questionnaire survey for nearly 5000 students conducted by the Family Education Research Institute of China Youth & Children Research Center in 2011 in Beijing, Guangzhou, Nanjing, Harbin, Shijiazhuang, Xi'an, Chengdu and Yinchuan, the payment for private tutoring has constituted a major part of families' education expenditure: 76% of the families spend an average of RMB 3, 820 on private tutoring each year, with the highest hitting RMB 80, 000. For urban families, children's education costs account for 76.1% of their total cost, 35.1% of the family's total expenditure and 30.1% of the family's total income (Guangming Daily, 2012-02-27). In Japan according to the statistics of the Ministry of Education, a child needs a total education expenditure of 10 million yen (about RMB 800, 000) from elementary through to tertiary education, most of which is spent on private supplementary tutoring (Sina Education, 2012-07-18). A sampling survey made by Korean education authorities reveals that in 2007 the country's elementary, junior high and senior high school students attended 10.2 hours of private tutorial classes per person per week on average, which cost KRW 288, 000 per person per month or KRW 3.46 million (about USD 3, 460) per person per year (Xinmin Weekly, 2010-08-04).

Since shadow education not only concerns the studies and daily life of students but also produces a certain impact on families' education expenditure and social stratification, it has drawn wide attention from the government, policy makers and scholars (Bray et al. 2014). It is, therefore, important and necessary to study the causal effects of private supplementary tutoring on students' learning performance so that parents can better understand such effects, rationally choose the strategy of supplementary tutoring and improve the return on investment. Studying the causal effects of private supplementary tutoring can also provide a useful reference for the government to allocate education resources more efficiently and more fairly. Due to the high cost of supplementary tutoring, the percentage of students from low-income families who receive it may be smaller than that of students from high-income families, thus leading to unequal opportunities of extracurricular education. But, will such unequal opportunities expand the inequality of educational outcomes? The paper tries to answer this question by analyzing the causal effects of shadow education on the learning performance of Chinese, Japanese and Korean students, based on the statistics of PISA 2012 conducted for students in Shanghai, Hong Kong, Japan and Korea.

The subsequent text is organized as follows: The second section is a review of literature about the effect of supplementary tutoring on students' learning performance and about the social inequality that shadow education may cause. In this section, the authors will also point out the main questions the present study needs to answer and the research hypothesis that needs to be verified. The third section will introduce the basis of sample selection and describe the method of the econometric model and the definitions of variables. The fourth section will present the results of empirical analysis. First, based on descriptive statistics the paper will analyze the difference in the percentage of attending math tutoring between different groups of students as well as the difference in math scores between the students attending private supplementary math tutoring and those not attending private supplementary math tutoring. Second, the paper will estimate the causal effect of private tutoring by means of reweighting propensity score matching (RPSM); analyze the effects of the length of tutoring time (LTT) for different school courses on math scores using a two-level linear model; and verify whether shadow education will expand inequality of educational outcomes. The fifth section will draw a conclusion, have some discussions and make policy recommendations

LITERATURE REVIEW

When investigating the percentage of attending private supplementary tutoring, quite a number of researchers analyzed the factors affecting students' participation in private tutoring (Bray et al. 2014; Bray & Kwok 2003; Tansel & Bircan 2006). However, few researchers have studied the effect of private supplementary tutoring on students' learning performance (Zhang 2013). An early study by Stevenson and Baker, not only created the concept of shadow education but also analyzed its effect on Japanese high school students' college entry based on the data of a survey regarding private supplementary tutoring for high school students in Japan (Stevenson & Baker 1992). According to their study, private tutoring, by adopting appropriate teaching strategies, may satisfy the individual needs of students and help them to learn in a more efficient manner. Based on their studies in Japan, Sawada and Kobayashi found that a positive correlation exists between the amount of time students spend in attending private tutoring of math and their math scores (He Ruizhu et al. 2011). But according to a study conducted by Smyth in Ireland in 2008, there is no significant difference in the scores of final examination between students who receive private supplementary tutoring and those who do not. In his opinion, there are two reasons for this. One is that students receiving private tutoring are usually those who are already quite good at their studies yet still want to improve their scores, but it is very difficult for them to achieve the purpose due to the "threshold effect". The other is that the time spent in attending private tutoring is much less than that spent in attending formal education and examination scores cannot be increased significantly in a short period of time (Smyth 2008). This leads to some of the questions the paper tries to answer: In Shanghai or Hong Kong, will the effect of private tutoring be weakened due to the "threshold effect"? Will tutoring in a certain course be ineffective because only a very limited amount of time has been spent on it? Or will tutoring in a certain course occupy the time that should be used for the study of other courses and therefore adversely affect the student's scores of these courses?

Due to limitations of data and the endogenous nature, studies on the causal effects of private tutoring on students' learning performance are scarce. Using a multi-level linear model and a conditional quantile regression model, Zhang Yu comprehensively and thoroughly studied the causal effect of private supplementary tutoring on the college entrance examination scores of high school graduates in Jinan, China (Zhang 2013). According to her study, private tutoring on the whole does not have significant effects on the total score of college entrance examination for either urban or rural students, but significant positive effects are seen on low-capacity urban students and those from low banding urban schools. For some rural students, however, private tutoring even produces negative effects on the score of college entrance examination. Her studies also revealed that, with regard to college entrance examination, private tutoring not only produces quite different effects on the score of different courses and the total score, but also shows different effects on students from schools of different qualities. Using the method of instrumental variables, Suryadarma et al discovered in 2006 that private tutoring does not show significant causal effects on the scores of Indonesian fourth graders. In our present research we use the RPSM method to study the causal effect of private tutoring on the score of middle school students, and see whether it produces different effects on the learning performance of students from different countries/regions or from families of different economic and social backgrounds.

A lot of sociologists and educators focus their study on the problem of social inequality that shadow education may cause. Many of them worry that the existence of shadow education may threaten social fairness and become a mechanism for maintaining and expanding social inequality (Verdis, Athanasios 2002; Murawska, Barbara & Putkiewicz 2006; Smyth 2009). According to the latest research by Bray M. et al, the market of private supplementary tutoring has regenerated social inequality in the mainstream education system. Family income is a major factor influencing Hong Kong students between Grade 9 and Grade 12 when it comes to whether or not to take private tutoring; compared with students from elite government schools, fee-charging ESF schools and DSS schools, those from low banding schools and fee-free aided schools have much less opportunities to receive private tutoring (Bray M. et. Al., 2014)[v]. Studies by Lei Wanpeng (2005) and Xue Haiping et al (Xue Haiping and Ding Xiaohao 2009) also show that the opportunity for a student to receive private tutoring is closely related with the ESCS of his/her family: the higher the ESCS, the more the opportunities. Although there have been no empirical studies on the rate of personal economic returns that shadow education can generate, the human capital theory has already pointed out that education, as an important way of human capital investment, can bring a person good and lifelong economic benefits. For this reason, shadow education can become a mechanism for increasing human capital reserves. From this point of view, shadow education is very likely to maintain or even expand social inequality.

Since the test result of PISA 2012 was published at the end of 2013, the performance of Asian students has attracted worldwide attention, including the attention of foreign media. The Washington Post, for example, reported the test result on December 5, 2013 with an article entitled “Diligent Asian Students Dominate Global Exam”. PISA 2012 tested about 510, 000 fifteen-year-old students from 65 countries and regions around the world, with students from Shanghai China ranking first in math, language and science literacy tests. Excellent results were also achieved by students from China Hong Kong, China Taiwan, Singapore, Japan and Korea[In most of the countries/regions that ranked high in PISA 2009 and PISA 2012, the percentage of students receiving private supplementary tutoring was big, indicating that their high PISA test scores were attributed to not only formal education but also private supplementary tutoring.]. American students did not enter Top 20 in any course tested.

How to understand and explain the excellent PISA scores of Asian students is also one of the tasks of this paper. Through a comparative analysis of private supplementary tutoring received by students of Shanghai, Hong Kong, Japan and Korea, we hope to reveal the characteristics of different groups of students who receive shadow education, see whether shadow education can really improve the learning performance of students and whether it has any heterogeneity, and estimate the effect of LTT on the scores of students. We will also discuss the question of whether shadow education may expand inequality of educational outcomes so as to provide an empirical basis for the government to formulate policies for regulating and controlling shadow education. Using the PISA 2012 data for Shanghai, Hong Kong, Japanese and Korean students, we attempt to examine the following four research hypotheses:

Hypothesis 1: The probability of receiving shadow education is higher for students from high-ESCS families than for those from low-ESCS families.
Hypothesis 2: Attending private supplementary math tutoring may significantly improve the math scores of students.
Hypothesis 3: Increasing the time of math tutoring may significantly improve math scores.
Hypothesis 4: Shadow education may aggravate the inequality of educational outcomes.

METHODS

Data Resources

Data used in the present research come from the official website of PISA (http://www.pisa.oecd.org) which is an internationally authoritative program for assessing the learning performance of students. The discipline literacy test and background questionnaire of PISA are designed by international disciplinary and measurement experts. Student discipline literacy is estimated using modern educational measurement theories (such as the item response theory) and professional statistical software (such as Conquest). The main object of assessment for PISA 2012 is mathematical literacy of students, which reflects not only their mastery of basic mathematical knowledge (about space & graphics; change & relationships; quantity; and uncertainty) but also their mathematical cognitive ability (the ability of mathematical expression, application and interpretation). The mathematical literacy scores (“math scores” for short hereinafter) of students are more comprehensive and representative than other scores.

Sample Selection

The PISA 2012 data for China Shanghai, China Hong Kong, Japan and Korea are used in the present research. The reasons why such data are used are that, first, the scale of shadow education is large in these countries/regions where people are deeply influenced by Confucianism, as is mentioned above; and second, these countries/regions outperformed others in three literacy tests of PISA 2012, and this paper is to explain to what an extent shadow education contributes to this excellent performance.

PISA 2012 tests were taken by 21, 231 Shanghai, Hong Kong, Japanese and Korean students aged from 15 years 3 months to 16 years 2 months, of whom 5, 177 came from 155 schools in Shanghai, 4, 670 from 148 schools in Hong Kong, 6, 351 from 191 schools in Japan and 5, 033 from 156 schools in Korea. Sampling was made using the method of two-stage probability proportional to size (PPS), with due consideration given to special-size schools (very small schools and very large schools). Therefore, the number of schools and the number of students sampled were different from country/region to country/region, but the two numbers should be more than 150 and 4500 respectively, so as to ensure the accuracy of estimated inter-school variance and intra-school variance. Seen from the above data, Shanghai, Japan and Korea met and Hong Kong basically met the sampling requirement on the number of schools and the number of students.

Variable Selection and Econometric Model

Based on the nested relations constituted by the data of PISA 2012 student questionnaire and school questionnaire, the paper tries to evaluate the effect of private supplementary tutoring on student math scores. Although the addition of control variables to the linear regression model can make the regression simulate experimental results, it is difficult to ensure that suitable control variables can be found to make conditional independence assumption (CIA). CIA is the core assumption that can give regression a causal explanation. Sometime it also means that selectivity bias comes from observable variables. In other words, selectivity bias will disappear when observable variables are introduced. validly established. Even if some control variables for student level and school level are introduced into the present research, it is still difficult to verify the validity of CIA. In the past two decades, matching became a tool for empirical research and attracted the interest of econometrists. The attractiveness of the matching strategy is that after some kind of propensity score matching both the experimental group and the control group are very close to the result of random assignment, which ensures that after a series of covariates have been brought under control the experimental result and the treated assignment become independent from each other, meaning that CIA is valid. By now we can get a causal explanation of the regression coefficient. The present research adopts the RPSM method proposed by Rosenbaum and Rubin in 1983 to estimate the causal effect of private tutoring on student learning performance. By this method, re-sampling is made through bootstrapping to obtain the standard error of the estimate of treatment effect so as to make up for the defect of the previous PSM method by which the standard error of treatment effect cannot be obtained. According to this method, the average treatment effect (ATE), average treatment effect treated (ATET) and average treatment effect non-treated (ATENT) are estimated via the following formulas (Giovanni Cerulli 2012):

$$\widehat{\text{ATE}}=\frac{1}{N}\sum_{i=1}^{N}\frac{[w_i-\hat{p}(\mathbf{x}_i)]y_i}{\hat{p}(\mathbf{x}_i)[1-\hat{p}(\mathbf{x}_i)]} \quad (1)$$

$$\widehat{\text{ATET}}=\frac{1}{N}\sum_{i=1}^{N}\frac{[w_i-\hat{p}(\mathbf{x}_i)]y_i}{\hat{p}(w=1)[1-\hat{p}_i(\mathbf{x})]} \quad (2)$$

$$\widehat{\text{ATENT}}=\frac{1}{N}\sum_{i=1}^{N}\frac{[w_i-\hat{p}_i(\mathbf{x}_i)]y_i}{\hat{p}(w=0)\hat{p}(\mathbf{x}_i)} \quad (3)$$

Where, “w_i” is a dummy variable indicating whether or not treatment is received (In this paper, it indicates whether or not math tutoring is received). If yes, w_i=1; if no, w_i=0. “$\hat{p}(x)$” is the estimated value of E (w|x). “y” is the outcome variable (In this paper, it represents math scores). “N” is sample size.

In addition to revealing the causal effect of private tutoring on learning performance using the RPSM method, we will also give the estimates of tutoring effects of the multi-level linear regression model after the same control variables are introduced in the robustness test, and compare these estimates with those obtained from the RPSM method.

The dependent variable in the present research is student math score. Why not choose reading score or science score as dependent variable of the econometric model？This is because the main object of assessment for PISA

2012 is mathematics, and the quantity of problems for reading and science literacy tests is not enough for comprehensively assessing student reading and science literacy. It should be noted that in the source data of PISA 2012, five plausible values (PVs) are provided. PVs are five values randomly drawn from a posterior distribution of student ability obtained based on student response, background variables and the marginal estimation technique of the item response theory. Intuitively speaking, they are values representing the scope of ability that a student may have. for each of the reading, math and science literacy tests. Generally speaking, since PVs have a certain degree of randomness, it is not suitable to report them as individual scores to the students. But they have irreplaceable advantages in estimating the population parameters. For instance, PVs can be used to obtain the unbiased estimates of population parameters, and in a complex sampling design they can be used to obtain the accurate estimates of standard errors.

By referring to the practice of PISA data analysis, in the present research we simultaneously use the 5 PVs in the multi-level model. Dependent variables are divided into two layers. The first layer consists of student-level variables including student gender, family ESCS and private supplementary math tutoring. The second layer consists of school-level variables including school type (public or private) and school location (urban or rural).

$$\begin{aligned} MS &= \alpha_0 + \alpha_1 gender + \alpha_2 escs + \alpha_3 mathoutschool + \varepsilon \\ \alpha_0 &= \gamma_{00} + \gamma_{01} public + \gamma_{02} location + \mu_0 \end{aligned} \tag{4}$$

In the above formulas, MS represents student math score. The meanings of other variables are listed in Table 2. When setting the education production function model for analyzing the variation of student learning performance, gender, family ESCS and school type are usually taken as control variables. This way, experimental results can be simulated even in the absence of random assignment and therefore, the regression coefficient of key explanatory variables can be regarded as an approximate estimate of the causal effect (Angrist & Pischke 2008). These variables are chosen by referring to the findings of related studies. According to the studies of Lei Wanpeng (2005) and Xue Haiping et al (2009), important factors influencing student access to private supplementary tutoring include family income, parents' education background, student learning performance, urban-rural disparity and regional disparity: students from big cities, public schools and higher-ESCS families are more likely to receive private supplementary tutoring than those from rural areas, private schools and low-ESCS families. In regard to the influence of student learning performance, however, the two scholars draw opposition conclusions. Xue Haiping's study (2009) reveals that students with good learning performance are more likely to attend supplementary tutoring, while Lei Wanpeng's research (2005) shows that students with poor learning performance are more likely to attend supplementary tutoring. According to a survey for Hong Kong secondary school students conducted by Bray M. et al (2004), girls spend much more money on private supplementary tutoring than boys, and the level of family income has a significant positive effect on private supplementary tutoring. A recent survey in Korea also shows that student access to private tutoring is directly related with parents' education background and family income level. Ninety percent of parents with a higher education background let their children attend private tutoring, while this percentage is only 50% for parents with a high school education background. Families with a monthly income of 7 million Korean won spend nearly10 times as much as those with a monthly income of 1 million Korean won on private tutoring for their children (China Education Daily, 2011-08-16). Therefore, these variables are also incorporated into the econometric model of the present study as independent variables that influence students' opportunity of receiving private tutoring. This is because when choosing the independent variables in the propensity score matching model we need to consider not only the factors that affect outcome variables but also those that affect the received treatment. So, the present study takes student gender, family ESCS, school type and school location as the covariates of the RPSM model in order that the estimates of tutoring effect are more approximate to the causal effect. In the present study, analysis results of the multi-level model are used as part of the robustness test of the results of the propensity score matching model. Therefore, the model is chosen mainly based on how the independent variables of the propensity score matching model are chosen. We use the same variables as those of the model to serve as the independent variables in the multi-level model so that the results of the two models are comparable.

Table 2: Names and Coding of Variables

Variables	Definition	Scoring
Student-level variables		
Gender	Student gender	Discrete variable, two-point scoring: 1-female, 0-male
ESCS	A combination of three indicators: highest occupational status of parents; highest education level of parents; the amount of family assets	Continuous variable
Mathoutschool	Whether or not to attend supplementary math tutoring	Discrete variable, two-point scoring: 1-

	on a weekly basis	Yes, 0-No
Langoutschool	Whether or not to attend supplementary language tutoring on a weekly basis	Discrete variable, two-point scoring: 1-Yes, 0-No
Scioutschool	Whether or not to attend supplementary science tutoring on a weekly basis	Discrete variable, two-point scoring: 1-Yes, 0-No
Otheroutschool	Whether or not to attend supplementary other courses tutoring not listed in the above three tutoring variables on a weekly basis	Discrete variable, two-point scoring: 1-Yes, 0-No
Mathtime	The amount of time spent attending math tutoring	The variable is treated as 0 hour if the student does not attend any private tutoring at all; as 1 hour if the student spends less than two hours attending private tutoring per week; as 3 hours if the student spends 2-4 hours per week; as 5 hours if the student spends 4-6 hours per week; as 6 hours if the student spends 6 or more than 6 hours per week.
Langtime	The amount of time spent attending language tutoring	
Scitime	The amount of time spent attending science tutoring	
Othertime	The amount of time spent attending tutoring in other courses	
School-level variables		
Classsize	Number of students in the class	Continuous variable
Public(school type)	The school is public or private	Discrete variable, two-point scoring: 1-public, 0-private
Location [a]	The school is in the city or in the country	Discrete variable, two-point scoring: 1- in the city，0-in the country

Notes: For PISA samples, school location is categorized into 5 types: village, small town, town, city and metropolis. In the present study, the first two types are classified as "the country", and the last three types are classified as "the city".

RESULTS and FINDINGS

Descriptive Statistics of Student-level and School-level Variables

Descriptive statistics of student-level and school-level variables of Shanghai, Hong Kong, Japan and Korea are listed in Table 3, which shows the following information. (1) In terms of student-level variables, Shanghai students have the highest math scores, followed by Hong Kong, Korean and Japanese students successively, and the math scores of students of the four countries/regions are all higher than the average math score (494) of students of OCED countries. In Shanghai, more girl students participated in the PISA test than boy students, but the situation was opposite in the other countries/regions. Korean and Japanese students had a higher average value of family ESCS than Shanghai and Hong Kong students. In Shanghai and Japan, more than 70% of students attend private supplementary math tutoring and 50% - 60% of students attend language, science and other tutoring. In Korea and Hong Kong, not so many students attend private supplementary tutoring, but math tutoring is attended by more students than other tutoring. In the four countries/regions, students' weekly mathtime is more than othertime, and the average weekly mathtime of Korean and Shanghai students (2.56 and 2.01 hours respectively) is longer than that of Japanese students (1.74 hours) and Hong Kong students (1.07 hours). (2) In terms of school-level variables, as many as 91% of the schools sampled from Shanghai are public ones, and this percentage for Japan and Korea is 70% and 53% respectively, but for Hong Kong it is as low as 7%, which is largely is largely due to the characteristics of Hong Kong's education system[vi].The average classsize of Shanghai and Japan is 39.29 and 36.29 respectively, slightly bigger than that of Hong Kong (34.08) and that of Korea (33.6). All the schools sampled from Shanghai and Hong Kong, 98.4% of the schools sampled from Japan and 93.7% of those sampled from Korea are located in the city, with rural schools accounting for only a small proportion.

Table 3: Descriptive Statistics of Variables (Mean /Standard Error)

	China Shanghai	China Hong Kong	Japan	Korea
student-level variables				
Math score	612.68(100.979)	561.24(96.31)	536.41(93.524)	553.77(99.077)
Proportion of girl students gender	0.51(0.500)	0.46(0.499)	0.47(0.499)	0.47(0.499)
ESCS	-0.36(0.964)	-0.79(0.973)	-0.07(0.713)	0.012(0.743)
Mathoutschool	0.71(0.455)	0.47(0.499)	0.70(0.459)	0.66(0.474)
Langoutschool	0.51(0.500)	0.24(0.430)	0.58(0.494)	0.53(0.499)
Scioutschool	0.55(0.497)	0.29(0.454)	0.54(0.498)	0.39(0.488)
Otheroutschool	0.57(0.495)	0.42(0.493)	0.69(0.461)	0.65(0.478)
Mathtime	2.01(1.905)	1.07(1.527)	1.74(1.814)	2.56(1.370)
Langtime	1.33(1.786)	0.44(1.019)	1.04(1.323)	1.99(1.108)
Scitime	1.49(1.820)	0.68(1.378)	0.95(1.279)	1.70(1.007)

Othertime	1.41(1.736)	0.90(1.438)	1.62(1.758)	2.37(1.281)
School-level variables				
Public	0.91(0.290)	0.07(0.251)	0.70(0.458)	0.53(0.499)
Classsize	39.29(7.558)	34.08(5.018)	36.29(5.662)	33.60(6.041)
Location [a]	1.00(0)	1.00(0)	0.98(0.127)	0.94(0.244)

Note a: For PISA samples, school location is categorized into 5 types: village, small town, town, city and metropolis. In the present study, the first two types are classified as "the country", and the last three types are classified as "the city".

Analysis of the Proportions of Different Groups of Students Attending Math Tutoring

As shown in Table 4, in all the countries/regions the percentage of attending math tutoring is bigger for students from high-ESCS families than those from low-ESCS families, which is a proof of Research Hypothesis 1. The gap is respectively 9.7 and 12.8 percentage points in Shanghai and Korea, and even as big as 17 percentage points in Hong Kong and Japan. Although the PISA questionnaire does not involve tutoring cost, the results of surveys in China, Japan and Korea, as mentioned above, show that private tutoring cost accounts for a large part of household education expenditure, which is unaffordable for low-income families. So, it is no wonder that in the four countries/regions the percentage of attending private tutoring is bigger for students from high-ESCS families than those from low-ESCS families. Will this difference in turn lead to a bigger gap in learning performance between the two groups of students? In other words, will private supplementary tutoring aggravate the inequality of educational outcomes? This is a question to be discussed in the text that follows.

In the three countries/regions except Hong Kong, the percentage of attending math tutoring is obviously higher for students of private schools than those of public schools, and this is largely because students of private schools are from relatively high-income families that can afford high tuition fees. In Hong Kong however, the percentage of attending private tutoring is higher for students of public schools than those of private schools. Since Hong Kong's per capita income is quite high[vii], families of both private and public school students can afford to pay private tutoring cost[viii]. In Hong Kong, private schools already provide high-quality education services so their students do not need to receive private supplementary tutoring. In Japan and Korea, the percentage of attending math tutoring is significantly higher for urban students than for rural students, although quite a proportion (58.5%) of Korean rural students attends math tutoring as well. In 2005 the Korean government launched an "After-school Education Program", according to which public schools shall provide after-school services such as care, custody, extra guidance to classroom lessons, art courses and recreational activities for their students. With central and local governments providing special funds for the program, universities and research institutes are responsible for developing specific projects of after-school education. With the financial support of the government, not only can urban students satisfy their individualized learning needs at a below-market price, but students from rural and low-income families can enjoy free after-school education, which narrows the education gap between students from high-income families and those from low-income families. In recent years, the Korean government has continuously strengthened the implementation of and investment in the "After-school Education Program". According to statistics of the Korean Ministry of Education, Science, and Technology, 99.8% of primary and secondary schools in Korea have started after-school education, benefiting nearly half of their students (China Education Daily, 2011-08-16).

Table 4: Percentage of Attending Math Tutoring and Difference Test for Different Groups of Students

		Shanghai		Hong Kong		Japan		Korea	
		Percentage	χ^2 test	Percentage	χ^2 test	Percentage	χ^2 test	Percentage	χ^2 test
SES[ix]	High	75.4%	648.453***	55.6%	1332.020***	78.2%	24210.139***	72.1%	7151.200***
	Low	65.7%		38.6%		61.5%		59.3%	
School Type	Public	70.1%	88.683***	53.7%	68.363***	67.8%	3263.822***	63.5%	1206.743***
	Private	76.3%		46.2%		74.5%		68.8%	
Gender	Boys	67%	344.901***	45.3%	46.746***	67.2%	2560.520***	65.4%	74.605***
	Girls	74.1%		48.4%		72.6%		66.7%	
School Location	Urban	70.7%		46.7%		70.4%	7878.170***	66.5%	652.008***
	Rural	——		——		32%		58.5%	

Note: "*" indicates a significant difference from zero at the 0.1 level; "**" indicates a significant difference from zero at the 0.05 level; and "***" indicates a significant difference from zero at the 0.01 level; the same below.

The above significant differences in the percentage of attending math tutoring between different groups of students will be further examined in the Logit Model in the text that follows. The analysis results here are also the basis for incorporating into the Logit Model SES, gender, school type, and school location as explanatory variables.

Learning Performance Differences between Students Who Receive Private Tutoring and Those Not[x]

By gender as shown in Table 5, in Shanghai the math score of boy students who receive private tutoring is roughly 22 points higher than that of those who do not and this difference is approximately 17 points for girl students. The difference in girl students is about 4.5 points smaller than that in boy students, showing no significant difference. Among Shanghai students who attend math tutoring, the math score of girls is significantly lower than that of boys (by 9.7 points), while among Shanghai students who do not attend math tutoring, the math score of girls is about 5.2 points lower than that of boys, showing no significant difference. Evidently, the difference in math scores does not totally result from tutoring. Our basic inference is that in shanghai math tutoring brings roughly the same academic benefits to boy students as it does to girl students and that private tutoring does not aggravate the difference in math scores between boys and girls. In Japan, the math score of boy students who receive private tutoring is 60 points higher than that of those who do not receive private tutoring, and for girl students this gap is 65 points, only 5 points more than that of boys, showing no significant difference, either. Among Japanese students who attend math tutoring, girl students have their math score 19 points lower than that of boy students, while among Japanese students who do not attend math tutoring, girl students have their math score 24.3 points lower than that of boy students, showing a very significant difference. Our inference is that in Japan math tutoring brings roughly the same academic benefits to boy students as it does to girl students but contrary to the case of Shanghai, private tutoring may narrow the difference in math scores between boys and girls.

In Shanghai, seen from the perspective of family ESCS, the math score of students from families below the average ESCS value (hereafter referred to as "low-SES" students) who attend math tutoring is 20.3 points higher than the math score of the other low-SES students who do not attend math tutoring; but the math score of students from families above the average ESCS value (hereafter referred to as "high-SES" students) who attend math tutoring is only 1 point higher than the math score of the other high-SES students who do not attend math tutoring. So a significant difference exists between the two types of students in terms of whether or not to receive private tutoring. Among the Shanghai students who attend math tutoring, high-SES students have their math score 60 points higher than that of low-SES students; among those who do not attend math tutoring, high-SES students have their math score 79.4 points higher than that of low-SES students. Therefore, our inference is that math tutoring brings more academic benefits to low-SES students than it does to high-SES students in Shanghai, and may narrow the learning performance gap between the two groups of students. We will verify this hypothesis in the subsequent text by means of an econometric model.

In Japan, high-SES students who attend math tutoring have their math score 58.4 points higher than that of those who do not, and low-SES students who attend math tutoring have their math score 48.2 points higher than that of those who do not. But there is no significant difference between the two types of students in terms of whether or not to receive private tutoring. Among the Japanese students who attend math tutoring, high-SES students have their math score 44.2 points higher than that of low-SES students, and among those who do not attend math tutoring, high-SES students have their math score 34.1 points higher than that of low-SES students. Therefore, our inference is that math tutoring brings more academic benefits to high-SES students than it does to low-SES students, and may broaden the learning performance gap between the two groups of students. This hypothesis will also be verified in the subsequent text by means of an econometric model.

The definitions and scoring of the independent variables used in the model and those used in the descriptive analysis in the subsequent text are listed in Table 2.

Table 5: Math Score Differences between Different Groups of Shanghai and Japanese Students in Terms of Whether or Not to Attend Math Tutoring

<table>
<tr><th colspan="2"></th><th>Attending</th><th>Not attending</th><th>Score difference between students who attend math tutoring and those who do not</th><th>Score difference between different groups of students in terms of whether or not attending math tutoring</th></tr>
<tr><td colspan="6">Shanghai</td></tr>
<tr><td rowspan="3">Gender</td><td>Girl</td><td>614.446
(93.476)</td><td>596.942
(109.792)</td><td rowspan="2">17.503***
(t=2.807)</td><td rowspan="3">-4.528
(t=-0.51)</td></tr>
<tr><td>Sample size</td><td>1312</td><td>456</td></tr>
<tr><td>Boy</td><td>624.164
(97.651)</td><td>602.132
(112.588)</td><td>22.032***
(t=3.040)</td></tr>
</table>

	Sample size	1123	545		
	Girl -Boy	-9.718** (t=2.446)	-5.189 (t=0.681)		
SES	High	646.010 (88.178)	645.040 (104.066)	0.97 (0.142)	-19.348** (t=-2.23)
	Sample size	1320	422		
	Low	585.946 (93.798)	565.628 (104.276)	20.319*** (t=3.319)	
	Sample size	1112	579		
	High - Low	60.064*** (t=10.390)	79.413*** (t=8.591)		
Japan					
Gender	Girl	547.875 (84.887)	482.687 (79.464)	65.188*** (t=10.617)	5.174 (t=0.68)
	Sample size	1463	564		
	Boy	566.981 (89.086)	506.967 (94.698)	60.014*** (t=11.628)	
	Sample size	1438	694		
	Girl -Boy	-19.107*** (t=-4.091)	-24.280*** (t=-3.778)		
SES	High	577.098 (83.885)	518.714 (90.031)	58.383*** (t=9.990)	10.136 (t=1.48)
	Sample size	1605	452		
	Low	532.896 (85.004)	484.648 (85.978)	48.248*** (t=10.155)	
	Sample size	1266	784		
	High - Low	44.202*** (t=8.620)	34.067*** (t=4.905)		

An RPSM-based Analysis of Tutoring Effects

As mentioned above, one of the advantages of matching strategy is that after some sort of propensity score matching, the experimental group and the control group both approach the results of random assignment, thus ensuring the validness of CIA so as to obtain a causal explanation of the regression coefficient. The present study adopts the RPSM model to reveal the causal effect of math tutoring on student math score.

Table 6: RPSM-based Estimates of Math Tutoring Effects

	Treated group sample size	Control group sample size	ATET (standard error)	ATENT (standard error)	ATE (standard error)
Shanghai	2432	1001	3.105 (23.576)	7.679 (57.017)	4.439 (16.660)
Hong Kong	1427	1608	-11.281 (34.457)	-9.911 (23.635)	-10.555 (14.224)
Japan	2871	1236	50.529*** (5.983)	46.947*** (17.132)	49.451*** (4.194)
Korea	2168	1107	50.222*** (9.524)	49.550*** (17.894)	49.994*** (6.150)

According to the descriptive statistics, after matching, mean values of the characteristic variables of the treated group and control group show no significant differences and have passed LR tests. The P values of LR test statistics of student samples of Shanghai, Hong Kong, Japan and Korea are respectively 0.323, 0.395, 0.959 and 0.063, meeting the requirement of CIA, indicating that matching quality is fairly good and the result is credible.

As shown in Table 6, the ATE of math tutoring on Shanghai students is 4.4 points, and math tutoring brings those who attend math tutoring a net effect of 3.1 points, 4.5 points lower than the tutoring effects on the students who do not attend math tutoring, indicating that students who do not attend math tutoring may benefit more from tutoring. However, the ATE of math tutoring is 49.5 points on Japanese students and 50 points on Korean students, and math tutoring brings Japanese and Korean students who attend math tutoring a net effect of 50.5 and 50.2 points respectively, 3.6 points and 0.7 point higher than the tutoring effects on the students who do not attend math tutoring. Obviously, Japanese and Korean students benefit much more from math tutoring than Shanghai students. This shows that Research Hypothesis 2 is valid for Japanese and Korean students who attend math tutoring but invalid for Shanghai and Hong Kong students who attend math tutoring. There are some negative values of tutoring effects for Hong Kong students, although none of the three types of effect values is statistically significant. As for the fact that tutoring effect varies greatly from country/region to country/region, we think this is due to two reasons. First, the quality of math tutoring is not high enough either in Shanghai or in Hong Kong, producing an insignificant effect on both “good” and “poor” students. Second, Shanghai students, especially high-SES students, have received very good formal education in school and their math scores are already high enough (see Table 5), leaving little room for further improvement. According to the “plateau phenomenon” in school education, high-SES Shanghai students have probably reached the plateau phase in their studies and therefore it is very difficult for them to raise their scores to a remarkable new height (what is called the “threshold effect”) . That means private supplementary tutoring can hardly bring significant additional benefits, a fact that frustrates the parents who send their children to tutoring classes. The present study does not see any obvious effect of math tutoring on Shanghai students. In Japan and Korea, however, since most students have not entered the plateau phase, they can learn a lot of knowledge, methods and skills by receiving private supplementary tutoring so as to make up for their knowledge loopholes and significantly improve their scores in the end. Supplementary tutoring can bring obvious benefits to Japanese and Korean students whether they have attended or are going to attend math tutoring. Therefore, it is rewarding and advisable for Japanese and Korean parents to send their children to tutoring classes. In Hong Kong, the percentages of students who attend math, science or language tutoring are relatively low, which is probably the result of rational decision-making of some Hong Kong parents. It is also noteworthy that in Shanghai tutoring effects are more significant on students who have not received any prior tutoring than those who have been attending private tutoring for a period of time. Most of the students who have not received private tutoring probably come from low-income families. As mentioned above, the proportion of high-SES students who receive private tutoring is nearly 10 percentage points higher than that of low-SES students who receive private tutoring (As shown in Table 7, family ESCS is one of the most important factors that determine whether parents send their children to tutoring classes or not). If low-SES students have the opportunity to receive private tutoring, they will be able to improve their learning performance.

Seen from Table 7, gender, family ESCS, school type and school location are all important factors that influence students’ attending or not attending math tutoring, a result consistent with that of Table 4; and in the four country/region models, the coefficient of the variable ESCS is significantly positive, once again proving Research Hypothesis 1: the probability of receiving shadow education is significantly higher for students from high-ESCS families than for those from low-ESCS families.

Table 7: The Logit Model Influencing Students’ Attending or Not Attending Math Tutoring

Variables	Shanghai	Hong Kong	Japan	Korea
Gender	0.324*** (1.382)	0.156** (1.169)	0.247*** (1.280)	0.087 (1.091)
ESCS	0.296*** (1.345)	0.410*** (1.507)	0.710*** (2.033)	0.532*** (1.702)
Location			1.223*** (3.397)	0.166 (1.181)
Public				-0.203*** (0.816)
Intercept	0.852*** (2.345)	0.135** (1.145)	-0.382 (0.682)	0.593*** (1.810)

Note: 1. Data not in parentheses are coefficients of the logit model; those in parentheses are odds ratios. 2. The variable of school type is not incorporated into the models of Shanghai, Hong Kong and Japan, because if it is incorporated it will be impossible to pass the balancing test of the treated group and control group after matching.

Robustness check

Given the same control variables, Table 8 shows a comparison of estimated math tutoring effects according to the OLS model and RPSM model. Seen from the table, the tutoring effects corresponding to the two methods in the four country/region models are the same in sign and significance, differing in value only. This means that the result of RPSM is basically reliable. The assumption of OLS for estimating tutoring effects is that whether a student receives tutoring or not is totally random. In reality, however, low-capacity students are more likely to attend math tutoring than high-capacity students (Lei Wanpeng 2005; Yu Zhang 2013) because the latter have the ability to improve their scores through self-study. So, the OLS model may have underestimated tutoring effects because it does not take into consideration the variable of student capacity that has a negative correlation with the likelihood of receiving private tutoring. As for why tutoring effects on Shanghai and Hong Kong students are so different from the tutoring effects on Japanese and Korean students, the text below will give a detailed analysis and explanation from the perspective of how LTT and tutoring in other courses affect the math score of students.

Table 8: Estimated Math Tutoring Effects According to the OLS model and RPSM Model When Given the Same Control Variables

	OLS	RPSM(ATET)
Shanghai	0.332(3.274)	3.105(23.576)
	[3433]	[2432/1001]
Hong Kong	-15.838***(2.826)	-11.281(34.457)
	[3035]	[1427/1608]
Japan	13.468***(2.964)	50.529***(5.983)
	[4107]	[2871/1236]
Korea	24.682***(3.366)	50.222***(9.524)
	[3275]	[2168/1107]

Note: 1. Control variables in the OLS model of the four countries/regions are the same as the independent variables in Table 7. 2. The data in parentheses "()" are standard errors; the data in the square brackets "[ ]" beneath the OLS results are sample sizes; and those in the square brackets "[ ]" beneath the RPSM results are sample sizes of the treated group/control group.

Effects of LTT on student scores

Why are math tutoring effects so different among Shanghai, Hong Kong, Japanese and Korean students? According to the fore-mentioned descriptive statistics, the percentage of students receiving supplementary tutoring varies significantly from course to course and from country/region to country/region, and LTT of students in each country/region is obviously different.

For instance, on average Hong Kong students spend less time receiving math and science tutoring than students of the other three countries/regions. Will tutoring classes for different school courses produce an additive effect or crowding-out effect? We will use a two-level linear model to study the differences in tutoring effects on the students of the four countries/regions based on the differences in LTT.

Table 9: Effects of LTT on Math Scores: Results of A Two-level Linear Model

	Model 1		Model 2	
Variables	Shanghai	Japan	Shanghai	Japan
Fixed effect				
Intercept	192.184***	330.885***	201.503***	329.177***
Student-level variables				
Gender	-14.148***	-16.556***	-14.799***	-16.569***
Escs	7.716***	5.385**	7.641***	5.166**
$Escs^2$	-5.805***	0.870	-5.655***	0.714
Mathtime	1.107	8.025***	-0.038	6.344***
$Mathtime^2$	-0.336	-0.593*	0.003	-0.524
Langtime			-5.301***	0.567
Scitime			3.485***	-0.687
Othertime			1.547	2.261***
School-level variables				
Public	-33.611***	-3.681***	-33.376***	-3.796***
Classsize	24.466***	7.843***	23.899***	7.958***
	-0.308***	-0.057***	-0.301***	-0.059
Random effect				
Intra-school	5172.398***	3876.623***	5122.177***	3864.359***
Inter-school	3431.134***	3280.015***	3237.169***	3272.459***
N	3433	4107	3425	4095

Table 10: Effects of LTT on Math Scores: Results of A Two-level Linear Model (continued)

	Model 1		Model 2	
Variables	Hong Kong	Korea	Hong Kong	Korea
Fixed effect				
Intercept	524.043***	741.693***	532.539***	735.406***

Student-level variables				
Gender	-23.166***	-10.015**	-22.267***	-10.737**
Escs	8.352***	17.646***	9.221***	17.402***
Escs^2	0.581	5.348**	1.106	5.364**
Mathtime	-7.697**	13.709***	-10.002***	13.766***
Mathtime^2	0.621	-0.562	0.876	-0.705
Langtime			-9.649***	0.359
Scitime			8.647***	-3.096
Othertime			0.792	3.804***
School-level variables				
Public	31.150***	-17.033***	29.688***	-15.710***
Classsize	-1.175	-16.154***	-1.615	-16.043***
Classsize^2	0.083**	0.286***	0.087*	0.285***
Random effect				
Intra-school	5142.270***	5713.179***	4962.542***	5686.631***
Inter-school	3207.368***	2596.605***	3026.882***	2619.061***
N	3035	3275	3016	3238

Note: "***" indicates significance at the 0.01 level; "**" indicates significance at the 0.05 level; and "*" indicates significance at the 0.1 level.

Comparing the results of Model 1 and Model 2 in Table 9, we see that differences in the coefficients of gender, ESCS, school type and classsize are very small, indicating good robustness of the models and high reliability of the research results. Here, we focus on the effects of LTT on student math scores. Seen from the results of Model 1 and Model 2, mathtime has a significant positive effect on student math scores, but such an effect is shown in an inverted U-shaped curve, according to which math tutoring effect reaches the maximum when weekly mathtime is around 7 hours (Currently, Japanese students have 1.74 hours of mathtime per week on average). Also seen from the results of Model 1 and Model 2, mathtime does not produce any significant nonlinear effect on the math scores of Korean students; its first-term effect is very significant; and unit mathtime may improve a student's math score by as many as 13.7 points. Thus, Research Hypothesis 3 that "increasing the time of math tutoring may significantly improve math scores" is verified in the cases of Japan and Korea. Significance (big value) of the effect of unit mathtime on Japanese and Korean students is also one of the reasons why the RPSM models give bigger ATT results for Japanese and Korean students than for Shanghai and Hong Kong students. By analyzing the sample of Shanghai students we find that mathtime has no significant effect, langtime has a significant negative effect and scitime has a significant positive effect on student math scores. These findings are consistent with the coefficients of scioutschool of Model 2 and Model 3 in Table 10, proving that science tutoring and math tutoring have a mutually additive effect. For Hong Kong students, mathtime shows an extremely negative effect on math scores, one of the reasons for which is probably that their weekly mathtime is too short (only 1.07 hours on average). Only when weekly mathtime is increased to at least 6 hours will it be possible to see substantial improvement in math scores. Another reason is probably that Hong Kong math tutors are not experienced enough, unable to teach students according to their aptitude or take targeted measures to improve their scores. Relatively speaking, science tutoring brings more benefits to Hong Kong students, with per unit of scitime improving math scores by as many as 8.6 points.

Does "shadow education" aggravate the inequality of educational outcomes?

Here in this part, we will discuss the question of whether shadow education aggravates the inequality of educational outcomes. For this purpose we use a linear regression model into which interaction terms are incorporated to see whether math tutoring produces different effects on the math score of students from different family backgrounds, in an attempt to provide an empirical basis for the government to formulate policies for regulating and controlling shadow education.

Table 11: Effects of Shadow Education on Math Scores: Results of A Two-level Linear Model

	Model 1		Model 2		Model 3	
Variables	Shanghai	Japan	Shanghai	Japan	Shanghai	Japan
Fixed effect						
Intercept	191.859***	326.862***	196.190***	325.535***	200.650***	325.810***
Student-level variables						
Gender	-14.161***	-16.119***	-15.151***	-16.009***	-15.152***	-16.008***
Escs	7.594***	5.539***	7.916***	5.498***	13.713***	6.923*
Escs^2	-5.842***	0.887	-5.798***	0.911	-5.319***	1.207
Mathoutschool	0.026	13.480***	4.224	8.481**	0.876	8.087**
Langoutschool			-11.959***	5.500*	-12.638***	5.475*

Scioutschool			4.438	2.260	4.625	2.271
Escs*mathoutschool					-7.774***	-1.985
School-level variables						
Public	-33.603***	-4.197***	-33.600***	-4.479***	-33.511***	-4.484***
Classsize	24.461***	8.064***	24.276***	8.113***	24.180***	8.110***
Classsize^2	-0.308***	-0.059***	-0.306***	-0.060***	-0.305***	-0.060***
Random effect						
Intra-school	5175.638***	3890.851***	5157.777***	3885.222***	5146.864***	3884.697***
Inter-school	3427.266***	3491.201***	3302.115***	3480.021***	3290.641***	3482.441***
N	3433	4107	3428	4104	3428	4104

Note: "***" indicates significance at the 0.01 level; "**" indicates significance at the 0.05 level; and "*" indicates significance at the 0.1 level.

Table 12: Effects of Shadow Education on Math Scores: Results of A Two-Level Linear Model (continued)

	Model 1		Model 2		Model 3	
Variables	Hong Kong	Korea	Hong Kong	Korea	Hong Kong	Korea
Fixed effect						
Intercept	518.973***	746.548***	524.957***	746.078***	524.028***	746.206***
Student-level variables						
Gender	-22.741***	-10.076**	-22.051***	-10.128**	-22.071***	-10.136**
Escs	8.118**	18.980***	8.780***	18.919***	11.062***	18.333***
Escs^2	0.466	5.728***	0.826	5.535**	1.167	5.439**
Mathoutschool	-15.777***	24.511***	-18.484***	23.588***	-21.434***	23.575***
Langoutschool			-17.490***	3.759	-17.749***	3.807
Scioutschool			21.544***	-2.207	21.571***	-2.200
Escs*mathoutschool					-3.924	0.882
School-level variables						
Public	30.847***	-17.300***	29.705***	-16.674***	29.640***	-16.688***
Classsize	-0.761	-15.669***	-1.063	-15.680***	-0.897	-15.692***
Classsize^2	0.076*	0.280***	0.079*	0.280***	0.076*	0.280***
Random effect						
Intra-school	5139.319***	5771.229***	5012.810***	5774.342***	5009.332***	5774.030***
Inter-school	3233.064***	2664.185***	3079.023***	2686.767***	3081.177***	2686.811***
N	3035	3275	3025	3248	3025	3248

First, looking at the coefficients of student-level independent variables we find that math scores of students of the four countries/regions vary significantly by gender, boys performing better than girls[xi]. For Japanese and Hong Kong students, family ESCS exerts a significant positive effect on math scores. For Shanghai students, however, the effect of family ESCS on math scores is shown in an inverted U-shaped curve: for the students whose family ESCS is below a certain value (0.649), the effect of ESCS on mathematical literacy increases as their family ESCS rises, but for those whose family ESCS is above the certain value, the effect of ESCS on mathematical literacy decreases as their family ESCS rises[xii]. Different from the case of Shanghai, the effect of family ESCS is presented in a U-shaped curve for Korean students. Math tutoring has a significant positive effect on Japanese and Korean students, a slight positive effect on Shanghai students but a significant negative effect on Hong Kong students. This result coincides with the fore-mentioned RPSM model result and the LTT model result. Language tutoring has a significant negative effect on the math scores of Shanghai and Hong Kong students[xiii], but a positive effect on the math scores of Japanese and Korean students, especially, a statistically significant positive effect on Japanese students. This is probably because the average langtime of Shanghai students is about 0.5 hour longer than that of Japanese students (as shown in Table 3), thus crowding out their mathtime. In addition, since language learning is very different from math learning in terms of method and way of thinking, it does not produce the same additive effect on math learning as science tutoring does. For Hong Kong students, science tutoring has a significant positive effect on math scores. For Shanghai and Japanese students, it also produces a positive effect on math scores although the effect is not statistically significant. This result also proves that math tutoring and science tutoring have a mutually additive effect.

Second, by looking at the coefficients of school-level independent variables, we find that school type has a significant effect on the math scores of students of all the four countries/regions. In Shanghai and Korea respectively, the math score of public-school students is 33 points and 17 points lower than that of private-school students, and this difference is 4 points in Japan. The case is just the opposite in Hong Kong, where the average math score of public-school students is 30 points higher than that of private-school students. Classsize exerts a significant inverted U-shaped effect on student math scores in Shanghai and Japan. In Shanghai when classsize reaches 40 students, it will produce the maximum positive effect on math scores, and in Japan this figure is 67 students. The average classsize in Shanghai is 39 students, already very close to the optimum scale, while in Japan there is still much room for increasing classsize. Unlike the situation is Shanghai and Japan, Hong Kong schools have some room for increasing classsize (the quadratic term of classsize is significantly positive but the first term is not so significantly positive), and in Korea a positive effect does not exist until classsize reaches 56 students, indicating that there is much room for classsize to increase.

Finally, by looking at coefficients of the interaction term of family ESCS and attending math tutoring in Model 3, we find that attending math tutoring may narrow the math score gap caused by different family ESCS. This effect is most significant in Shanghai where math score gap can be narrowed by about 8 points. In Hong Kong and Japan it can be narrowed by 4 and 2 points respectively. Coefficients of the Korean student learning performance model are positive, showing no statistical significance. Results of the Shanghai student learning performance model overturn Research Hypothesis 4 that "shadow education may aggravate the inequality of educational outcomes." According to the fore-mentioned descriptive statistics, in Shanghai math tutoring brings bigger benefits to low-SES students than to high-SES students. Therefore, shadow education may narrow the gap in student learning performance caused by the difference in family ESCS, thus promoting equality of educational outcomes. In Shanghai the average math score of low-SES students is 62.6 points lower than that of high-SES students, quite a big gap, and those low-SES students who have not received math tutoring have an average math score of a merely 565 points. If receiving math tutoring, they can markedly improve their scores because there is much room for improvement. The average math score of high-SES students who have not received math tutoring is already as high as 645 points, leaving little room for further improvement. For them, math tutoring is unlikely to bring significant benefits, or may even crowd out the time for them to develop their potentials in other aspects. Over the past two years China has made a lot of efforts to lighten the burden on primary and secondary students by shortening school time and substantially reducing the amount of homework, which has stimulated the enthusiasm of many parents and students to seek shadow education in an increasingly competitive way (Xue Haiping et al 2014). In this context low-SES parents may send children to tutoring classes so as to make up for their inability to academically instruct the children and narrow the score gap between their children and other students. The present study shows that in Shanghai it is financially advisable for low-SES parents to send their children to private supplementary tutoring classes but it is not so advisable for high-SES parents because the benefit is very limited. Seen from the coefficients of all the control variables in Models 1-3 shown in Table 10, the research results have good robustness and high reliability.

CONCLUSION

Based on the student sample data of PISA 2012 for Shanghai, Hong Kong, Japan and Korea and adopting an RPSM model and multi-level linear regression model, the present study has analyzed the effects of shadow education (private supplementary tutoring) on student math scores and come to the following conclusions:

1. Supplementary math tutoring may help Shanghai, Japanese and Korean students to improve their math scores, and such an effect is significantly positive for Japanese and Korean students. This conclusion is similar to the one drawn by OCED using PISA 2012 data (OECD 2013), although the measurement quantitative method used in the OCED PISA Report is different from the one used in our study. As mentioned above, Confucianism has a widespread influence in China, Japan and Korea, where the vast majority of parents want their children to go to high-quality universities so that they can have a good career in the future. They are willing to spend money on supplementary tutoring because, one the one hand, it helps children make progress in their studies and on the other hand, it enables children to achieve higher human capital which in the future will bring them higher returns on human capital investment. Numerous facts have shown that supplementary tutoring does enhance student performance, and therefore, parents are willing to pay high tuition fees in exchange for their children's advantage in competitive examinations.

2. Math tutoring effects are more significant on Japanese and Korean students than on Shanghai and Hong Kong students. We speculate that this is probably because the current overall mathematical literacy of Shanghai students is higher than that of Japanese and Korean students. The average math score of Shanghai students who receive math tutoring is already as high as 618.9 points, while this figure is only 557.3 for Japanese students and 575.4 for Korean students. Therefore, math tutoring in Shanghai is mainly intended for "top students" while that in Japan and Korean is mainly for "relatively poor students". Due to the "threshold effect" in learning, math tutoring creates less space for Shanghai students to further improve math scores than for Japanese and Korean students. In addition, the fact that per unit of mathtime generates a bigger tutoring effect for Japanese and Korean students is not only a reflection of the higher quality of math tutors of the two countries but may also be one of the reasons why the overall effect of math tutoring is more obvious for students of the two countries than

for Shanghai and Hong Kong students. In Hong Kong the effect of math tutoring is negative. This is probably because, on the one hand, weekly mathtime is too short (1.07 hours on average) to produce any substantial effect and on the other hand, private math tutors in Hong Kong do not have a thorough understanding of the curriculum content of formal schools, thus making tutorial efforts untargeted and tutoring effects insignificant. Another possible reason is that the RPSM model for measuring the tutoring effect on Hong Kong students missed the key characteristic variables that explain why students attend supplementary tutoring, thus making the result unrealistic. We also speculate that the formal school education system plays a major role and the shadow education system plays only a very limited role in helping Shanghai and Hong Kong students perform outstandingly in PISA 2012. For Japanese and Korean students who also performed excellently in PISA 2012, however, the role of shadow education is relatively bigger. That is why over the past years the Korean government has vigorously promoted investment in its “After-school Education Program”, a program that has benefited a great number of students.

3. Tutoring in different courses produces different effects on math scores: science tutoring and math tutoring have a mutually “additive effect”. In Shanghai, Hong Kong and Japan, tutoring in all science-related courses can improve student math scores. For Shanghai and Hong Kong students in particular, the effect of increasing the time length of science-related tutoring on math scores is significantly positive. This is mainly because science-related courses and mathematics have a lot in common, such as requiring logical reasoning and similar research paradigms. Language tutoring produces a significant negative effect on math scores in Shanghai and Hong Kong, but the case is just the opposite in Japan. We speculate that this is probably because the average weekly langtime of Shanghai students is longer than that of Japanese students, thus crowding out their mathtime, and that the weekly langtime of Hong Kong students not only is too short to produce any substantial effect but also reduces student mathtime, thus causing a decline in their math scores.

4. The opportunity of receiving private supplementary tutoring is obviously different for students with different family economic, social and cultural backgrounds. In the four countries/regions, the percentage of high-SES students who receive math tutoring is obviously higher than that of low-SES students: the gap is 10 percentage points in Shanghai and as big as 17 percentage points in Hong Kong and Japan. As has been previously described, the cost of attending supplementary tutoring is quite a big sum for low-income families either in China, Korea or in Japan and therefore, influences their decision whether or not to send their children to cram schools.

5. Attending supplementary math tutoring may narrow student performance gap caused by differences in family economic, social and cultural status (ESCS), thus promoting equality of educational outcomes. The learning performance of students from high-ESCS families is significantly higher than that of those from low-ESCS families. The gap in math scores between the two types of students is respectively 63, 41, 49 and 53 points in Shanghai, Hong Kong, Japan and Korea. By referring to the gender parity index developed by the UNESCO Institute for Statistics (UIS), family background parity indexes of all the four countries/regions we have calculated are bigger than 1.03, beyond the range of equality, although their gender parity indexes are within the range of equality (See attached Table 2). Therefore, it is necessary to take educational intervention measures to narrow the learning performance gap between students of different family backgrounds.

The big gap in learning performance mainly caused by differences in family background can be reduced by supplementary tutoring. In Shanghai for example, such a gap can be markedly narrowed by nearly 8 points. This conclusion may seem contrary to what people generally think. Most people regard private tutoring as an investment option of wealthy parents for their children, a thing that few ordinary parents are willing to do due to high costs. In their opinion, shadow education aggravates the inequality of educational outcomes or even social inequality. However, according to the data of PISA 2009 and PISA 2012 for Shanghai, around 70% students receive supplementary math tutoring, and this percentage is over 60% even for low-SES students. As mentioned afore, a survey conducted by the “Project Team for Research on Education Expenditure of Chinese Urban Families during the Compulsory Education Phase” reveals that 76% families choose to send their children to supplementary tutoring classes. It should be mentioned that according to the results of the RPSM model for Shanghai students, tutoring effects on the students who have not received any prior tutoring are more significant than on those who have been attending supplementary tutoring for a period of time. Most of the students who have not received any prior tutoring are very likely to come from low-income families, and as mentioned afore, their proportion of attending supplementary tutoring is almost 10 percentage points lower than that of high-income families (As shown in Table 7, family ESCS is an important factor influencing parents’ decision whether or not to let their children receive tutoring). If low-SES students receive supplementary tutoring, they will be able to significantly improve their learning performance and catch up with those high-SES students, thus making shadow education an “equalizer” of human capital accumulation. From this perspective, the government and education policy-makers should see the rational elements of shadow education and give it a “legal” status.

The present study leads to the findings that although the effect of supplementary math tutoring varies from country/region to country/region, it does improve student math scores and narrow the learning performance gap caused by differences in family ESCS, thus playing a certain role in promoting the equality of educational outcomes (In Shanghai, such an effect is significant). Quite a few researchers have found that math tutoring has become a must for a lot of low-income families in East Asian countries (Bray & Kwo 2013; Lee et al 2009; Lin & Chen 2006). According to surveys conducted in China, Japan and Korea, the cost of supplementary tutoring constitutes a considerable part of the total income of low-income families or even lowers their living standard. For this reason, we suggest that the government and schools provide free necessary tutorial services for low-performance students from these families and give full play to the role of school education as a "social equalizer"[xiv]. To this end, other countries/regions may follow the example of the Korean government to implement their own "after-school education programs", by which full-time teachers tutor low-performance students from low-income families after school, with all or part of the cost covered by the central or local government. It is also suggested that large-sized extracurricular education groups establish after-school learning funds and/or incentive funds[xv], for students from poor families, grant fee remissions to those from ultra-low income families and provide scholarships for poor students who have made substantial progress in their studies, so as to create a situation where the government, the school and the society jointly promote the equality of educational outcomes in the stage of compulsory education.

ACKNOWLEDGEMENT

An earlier version of this paper was presented on the CIES Annual Meeting organised by the College of Arts and Science and Cornell Higher Education Research Institute at the University of Washington, D.C. On March 8th-13th, 2015. Helpful comments were received from several scholars. The authors express appreciation to professor Rob Tierney from Sydney University for his helpful opinions and constructive suggestions. The research was supported by the Comprehensive Discipline Construction Fund of Faculty of Education at Beijing Normal University, and National Natural Science Foundation of China (No.71373165).

REFERENCES

Angrist, J.D. & Jorn-Steffen Pischke(2012). Mostly Harmless Econometrics: An Empiricist's Companion, MIT Press.

Baker D.P., LeTendre G.K. (2005). Demand for achievement: the worldwide growth of shadow education system in Baker D.P., LeTendre G.K. National Differences, *Global Similarities: World Culture and the Future of Schooling*. Stanford (Calif.): Standford University Press.

Bray, M., & Kwok, P. (2003). Demand for private supplementary tutoring: Conceptual considerations, and socio-economic patterns in Hong Kong. *Economics of Education Review*, 22 (6), 611–620

Bray, M. (2005). *Private Supplementary Tutoring: Comparative Perspectives on Patterns and Implications*. Sept. 13-15. Oxford: International Conference on Education and Development.

Bray, M. & Kwo, O. (2013). Behind the facade of fee-free education: Shadow education and its implications for social justice. *Oxford Review of Education*, 39 (4), 480–497.

Caritas, Community and Higher Education Service (2010). *Private Supplementary Tutoring of Secondary Students: Investigation Report*. Hong Kong: Caritas.

Mark Bray, Shengli Zhan, Chad Lykins, Dan Wang, Ora Kw (2014). Differentiated demand for private supplementary tutoring: Patterns and implications in Hong Kong secondary education. *Economics of Education Review* 38 , 24–37

Giovanni Cerulli (2012). Treatrew: a user-written STATA routine for estimating Average Treatment Effects by reweighting on propensity score. *RapportoTecnic*o, Anno 7, No.43, November.

He Ruizhu & Lu Naigui (2011). *Educational Quality and Equality Analysis for an International Perspective*. Educational Science Press, Beijing, 276 (In Chinese).

Japan, Ministry of Education and Training (2008). *Report on the Situation of Academic Learning Activities of Children*. Tokyo: MonbukagakushoHokokusho. (In Japanese).

Kim, Kyung-Keun (2010). "Educational Equality," in Lee, Chong Jae; Kim, Seongyuland Adams, Don (Eds.), *Sixty Years of Korean Education.* Seoul: Seoul National University Press, pp.285–325.

Lee, C. J., Park, H. J., & Lee, H. (2009). Shadow education systems. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of education policy research* (pp. 901–919). New York: Routledge for the American Educational Research Association.

Lei Wanpeng (2005). Supplementary Tutoring Expenditure on Senior High School Students: Influencing Factors and Policy Implications. *Education and Economy*, No.1, 39-42.(In Chinese).

Li, D., & Tsang, M. C. (2003). Household education decisions and implications for gender inequality in education in rural China. *China: An International Journal*, 1(2), 224–248.

Lin, D. S., & Chen, Y. F. (2006). Cram school attendance and college entrance exam scores of senior high school students in Taiwan. *Bulletin of Educational Research* (Taiwan), 52(4), 35–70, (in Chinese).

Murawska, Barbara & Putkiewicz, Elżbieta (2006): 'Poland', in Silova, Iveta; Būdienė, Virginija & Bray,Mark (eds.), *Education in a Hidden Marketplace: Monitoring of Private Tutoring*. New York: OpenSociety Institute, pp.257-277.

Ngai, Angela and Sharon Cheung (2010). *Students' Participation in Private Tuition*. Youth Poll Series No.188. Hong Kong: Hong Kong Federation of Youth Groups.

OECD (2013).PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices. Volume IV.Paris: OECD.

Ren Chunrong & Xin Tao (2013). A Follow-up Study on the Predicted Effect of Family Social and Economic Status on Primary Student Learning Performanc. *Education Research*, No.3, 79-87.(In Chinese)

Rosenbaum, P.R., and D.B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrica*, vol.70：41-55.

Sina Education. Asian Summer Holiday Unhappy: Shadow Education Also Prevailing in Thailand, Japan, Korea and India. 2012-07-18.(In Chinese). http://edu.sina.com.cn/kids/2012-07-18/113768276.shtml

Stevenson, David L. and David P. Baker (1992). Shadow education and Allocation in Formal Schooling Transition to University in Japan. *American Journal ofSociology*, 97(6):1639–1657.

Smyth, Emer (2008). The More the Better? Intensity of Involvement in Private Tuition and Examination Performance. *Educational Research and Evaluation*, 14 (5): 465-476.

Smyth, Emer (2009): Buying your Way into College? Private Tuition and the Transition to Higher Education in Ireland. *Oxford Review of Education*, 35 (1): 1-22.

Suryadarma, D., Suryahadi, A., Sumarto, S., & Rogers, F. H. (2006). Improving Student performance in public primary schools in developing countries: Evidence from Indonesia. *Education Economics*, 14(4), 401–429

Tansel, A., & Bircan, F. (2006). Demand for education in Turkey: A tobit analysis of private tutoring expenditures. *Economics of Education Review*,25(3), 303–313

Tsang, M. C. (2002). Comparing the costs of public and private schools in developing countries. In H. Levin & P. McEwan (Eds.), 2002 *Yearbook of the American Education Finance Association*. New York: Eye On Education.

Verdis, Athanasios (2002): *School Effectiveness Research for Educational Evaluation in Greece*. PhD thesis, Institute of Education, University of London.

Wang Qinghuan (2012). Private Tutoring Costs a Fortune for Ordinary Families. *Guangming Daily*. 2012-02-27.(In Chinese).

Xue Haiping & Ding Xiaohao (2009). A Study on Chinese Urban Student Supplementary Tutoring. *Education Research*, No.1, 39-46.(In Chinese).

Xue Haiping, Wang Dong & Wu Xiwei (2014). A Study on the Effect of Supplementary Tutoring on the Learning Performance of Left-behind Children in the Compulsory Education Stage. *Peking University Education Review*, vol.12, No.3, 50-62.(In Chinese).

Yu Zhang (2013). Does private tutoring improve students' National College Entrance Exam performance?—A case study from Jinan, China. *Economics of Education Review*, 32, 1–28.

Zhan Xiahong (2010). Various Types of Tutoring Form a Huge Business Chain in Korea: Elite Schools Compete Fiercely. *Xinmin Weekly*, 2010-8-4.(In Chinese).http://news.sina.com.cn/w/sd/2010-08-04/134920824548.shtml

Zhang Leisheng (2011). Korea: Extracurricular Tutoring "Fights to the End". *China Education Daily*, "International Education" section, August 16, 2011(In Chinese).

Berndt, T. J. (2002). Friendship quality and social development. *Current Directions in Psychological Science, 11*, 7-10.

Bernstein, M. (2002). 10 tips on writing the living Web. *A List Apart: For People Who Make Websites, 149*. Retrieved from http://www.alistapart.com/articles/writeliving

Biswas, S. (2008). *Dopamine D3 receptor: A neuroprotective treatment target in Parkinson's disease*. Retrieved from ProQuest Digital Dissertations. (AAT 3295214)

Brownlie, D. (2007). Toward effective poster presentations: An annotated bibliography. *European Journal of Marketing, 41*(11/12), 1245-1283. doi:10.1108/03090560710821161

Calfee, R. C., & Valencia, R. R. (1991). *APA guide to preparing manuscripts for journal publication*. Washington, DC: American Psychological Association.

Duncan, G. J., & Brooks-Gunn, J. (Eds.). (1997). *Consequences of growing up poor*. New York, NY: Russell Sage Foundation.

Kenneth, I. A. (2000). A Buddhist response to the nature of human rights. *Journal of Buddhist Ethics, 8*. Retrieved November 20, 2000 from http://www.cac.psu.edu/jbe/twocont.html

Plath, S. (2000). *The unabridged journals* K.V. Kukil, (Ed.). New York, NY: Anchor.

O'Neil, J. M., & Egan, J. (1992). Men's and women's gender role journeys: Metaphor for healing, transition, and transformation. In B. R. Wainrib (Ed.), *Gender issues across the life cycle* (pp. 107-123). New York, NY: Springer.

Schnase, J. L., & Cunnius, E. L. (Eds.). (1995). Proceedings from CSCL '95: *The First International Conference on Computer Support for Collaborative Learning*. Mahwah, NJ: Erlbaum.

Scruton, R. (1996). The eclipse of listening. *The New Criterion*, *15*(30), 5-13.

Wegener, D. T., & Petty, R. E. (1994). Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality & Social Psychology*, *66*, 1034-1048.

APPENDIX

Attached Table 1: Learning performance differences in Hong Kong and Korea between students who receive private tutoring and those who do not

		Attending	Not attending	Score difference between students who attend math tutoring and those who do not	Score difference between different groups of students in terms of whether or not attending math tutoring
Hong Kong					
Gender	Girl	553.662 (84.467)	554.047 (95.755)	-0.385 (t=-0.067)	-0.965 (t=-0.127)
	Sample size	681	723		
	Boy	570.909 (96.719)	570.329 (102.672)	0.58 (t=0.105)	
	Sample size	753	900		
	Girl -Boy	-17.247** (t=-2.221)	-16.282** (t=-2.385)		
SES	High	580.102 (89.134)	591.430 (96.475)	-11.328* (t=-1.920)	-6.931 (t=-0.909)
	Sample size	808	628		
	Low	540.199 (89.678)	544.596 (97.190)	-4.397 (-0.895)	
	Sample size	619	980		
	High - Low	39.903*** (t=6.066)	46.834*** (t=6.446)		
Korea					
Gender	Girl	560.404 (89.206)	513.631 (94.902)	46.773*** (t=7.100))	-28.664*** (t=-3.251)
	Sample size	1013	503		
	Boy	588.641 (95.652)	513.203 (101.713)	75.438*** (t=10.269)	
	Sample size	1157	606		
	Girl -Boy	-28.237*** (t=-4.193)	0.427 (t=0.051)		
SES	High	595.341	540.900	54.441***	-2.008

		(92.033)	(102.153)	(t=8.490)	(t=-0.239)
	Sample size	1237	476		
	Low	549.391 (89.333)	492.942 (90.849)	56.450*** (t=7.797)	
	Sample size	931	631		
	High - Low	45.950*** (t=8.368)	47.958*** (t=6.691)		

Attached Table 2: Gender parity indexes and family background parity indexes

	Average math scores of girls	Average math scores of boys	Gender parity indexes(GPI)[xvi]	Average math scores of high-SES students	Average math scores of low-SES students	Family background parity indexes[xvii]
Shanghai	609.888	615.607	0.991	643.246	580.588	1.108
Japan	527.011	544.884	0.967	563.015	514.120	1.095
Hong Kong	552.957	568.378	0.973	584.382	543.451	1.075
Korea	544.193	562.114	0.968	579.537	526.501	1.101

Notes

[i] According to Stevenson and Baker, when a change takes place in the courses of mainstream education, a corresponding change will take place in their “shadow courses”, and the development of mainstream education will lead to the expansion of shadow education (Stevenson & Baker1992).

[ii]. According to Mark Bray, “shadow education” does not include painting, piano playing, sports activities or other extracurricular activities that are not related to school courses.

iii. Private supplementary tutoring is a common phenomenon in Asian countries, but it is also increasingly seen in Europe, North America, Africa, Australia and other parts of the world.

[iv]. In the Confucian culture it is generally believed that a person’s success depends on constant efforts, and most parents hope that their children could improve their school performance by receiving private supplementary tutoring. In addition, deeply influenced by traditional belief that a child can make a success by working hard at their studies, parents attach great importance to their children’s education, including private supplementary tutoring..

[v] . In the school year of 2011-2012 Hong Kong had a total of 524 secondary schools, of which 497 were affiliated to the local education system and 27 were fee-charging ESF schools including those teaching in English. Of the 497 schools affiliated to the local education system, 32 were public schools directly owned and operated by the government and 365 were partly public schools funded or subsidized by the government. There were also 63 DSS schools covered by the Direct Subsidy Scheme and these schools were allowed to charge tuition fees and had some autonomy in curriculum setting and other aspects (Mark Bray et al, 2014

[vi]. According to the Education Bureau of Hong Kong, the SAR’s primary and secondary schools are classified into government schools, government funded/aided schools, DSS schools (“Direct Subsidy Scheme” schools), private schools, and fee-charging ESF schools. Government schools are schools established, operated and totally financed by the SAR government where all the teachers and administrative staff are civil servants. Of all the 1, 092 primary and secondary schools in Hong Kong in the academic year of 2011-2012, only 6% are government schools, or public schools.(http://www.edu.cn/xsc_12533/20130614/t20130614_963980.shtml)

[vii]. Data released by the World Bank shows that Hong Kong’s per capita GDP was USD 36, 708 in 2012, ranking the 27th place in the world.
http://data.worldbank.org/indicator/NY.GDP.PCAP.CD/countries

[viii]. Studies by Mark Bray, et al in 2014 reveal that monthly expenditure on private tutoring accounts for 8.7% of a family’s total monthly income on average in Hong Kong.

[ix]. Generated from ESCS, SES is a discrete variable with two-point scoring. “1” means being above the average value of ESCS of students of the participating country, and “0” means being below the average value.

[x]. Due to limited space, here we only discuss the learning performance differences between students who receive private tutoring and those who do not, by taking Shanghai and Japanese students as examples. For information in this regard about Hong Kong and Korean students, please refer to Attached Table 1 in the Appendix.

[xi]. Out of the 65 countries/regions involved in PISA 2012 tests, 37 see boy student math score significantly higher than that of girl students, and only 5 countries see the opposite situation. Similarly, among the 41 countries/regions involved in PISA 2003 tests, 27 see boy student math score significantly higher than that of girl students, 11 see boy student math score insignificantly higher than that of girl students, and only 1 country see boy student math score significantly lower than that of girl students.

[xii]. This result is basically consistent with the research result achieved by Ren Chunrong, et al in 2013 (Ren Chunrong & Xin Tao, 2013). This is probably because, on the one hand, the parents of high-SES students in China are usually busy with their work, thus having little time to share with their children and guide their study, or do not have high expectations on their children, and on the other hand, some high-SES students are not highly motivated to learn, making their learning performance inconsistent with the ESCS of their parents.

[xiii] .The average weekly langtime of Hong Kong students is only 0.44 hour, which cannot produce a substantial effect or may even crowd out their mathtime, causing a decline in math scores.]

[xiv]. School education is an open path by which people may break through their family origin and other restrictions, move up the social ladder and improve their lives. Just as the human capital theory stresses in particular, in a totally competitive market economy all people may increase their future income by investing in their own education, thus promoting economic fairness.

[xv] .By means of tax exemption or reduction, the government can encourage these large-sized extracurricular education groups to provide non-profit educational services for the poor.

[xvi]. Gender Parity Index (GPI) is an index designed by the UNESCO Institute For Statistics (UIS) to measure the relative access to education of males and females. In its simplest form, it is calculated as the quotient of the number of females by the number of males enrolled in a given stage of education. The more the quotient deviates from 1, the bigger the disparity. A quotient between 0.97 and 1.03 is considered in the parity range. This index is used in both 2010 and 2011 Global Education Digest

[xvii]. We calculate the family background parity index in a way that is similar to the way GPI is calculated.