ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

The Olympic Maths Educational Investment's Return

-- An Empirical Research Based on the Relationship between Olympic Maths Investment and High School Entrance Examination Score

Hongmei Li, Yihao Wang*

Faculty of Applied Economics, University of Chinese Academy of Social Sciences, Beijing, 102400, China

Abstract

Educational investment refers to the time and money devoted at present day, and the comparison with the future expected outcome. Olympic Maths studying is a popular kind of educational investment in China among primary school students, but its impact on future studies lacks statistical research. This thesis takes the Olympic Maths learning as an example, evaluates the impact of Olympic Maths investment on all subjects including Mathematics quantitatively. A high school Maths entrance examination score model is proposed in the thesis using Olympic Maths studying experience, along with other variables include parent education level, Maths anxiety level and Maths confidence level. A Propensity Score Matching (PSM) method is used to test the reliability of the model and the result turns out to be credible. The model suggests that the investment on Olympic Maths in primary school, along with high level of Maths confidence and parent education, low level of Maths anxiety will promote the Maths exam score significantly. The thesis later discussed the mechanism why these variables are effective.

Keywords

Educational Investment, Olympic Maths, High School Entrance Examination Score Model, Counterfactual Estimation.

1. Introduction and Literature Review

Investment decisions are focuses in the field of Finance. Investments are behaviors in which money is invested at present time and is expected to appreciate in the future. Traditional financial analysis studies the allocation of physical assets or financial assets to maximize profit when minimizing risks, and profound theoretical or empirical results have been worked out. On the contrary, educational investment behavior is never the focus of mainstream financial analysis.

If education itself is treated as a consumption good, it may be divided into public or private according to theories of public economics. The former includes compulsory education and higher education (There are disputes when treating the higher education as public good. Considering the exclusiveness of higher education, in this thesis it is treated as public good) while the latter contains personal education. Individual educational investments are acts taken by the educatees or his family members to take classes beyond the demand of compulsory education to gain future profits. Up to now, the importance and necessity of compulsory and higher education is widely discussed due to their popularity, no more discussion is needed in this thesis.

The profit of individual education behavior varies with the objects of educational investments, and studies on its profitability are needed. Due to their instincts of being private goods,

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

individual educational investments are similar to arranging physical assets or financial assets, which is to sacrifice the money and leisure time of present day in exchange for future return. According to rational person hypothesis, only when the net present value of future profit is greater than its cost, can the investment be regarded as appropriate. Meanwhile, when individual educational investments may bring the educatees extra gains, such as extra improvements in study or extra enhancement in ability or score, it is regarded highly profitable. Whether the decision should be made depends on the judgements of an educatee, which is to compare its cost with future gains.

There are plenty of references regarding personal educational investment theory. Yongxia Ma (2001) proved the adequacy of personal educational investments according to Human Capital Theory and Cost Compensation Theory. Jorgenson (1989) found that the cost paid by educational organizations are significantly lower than the human capital created in the same period when comparing the former with the wealth created by educated labors. Zhe Wu (2001) discussed the mechanism, function and status of educational investment theory in detail, and stated that educational investment should avoid falling into high consumption trap. Qiongxiu Xie et al. (2009) argued the profits of educational investments during all periods from three aspects, namely investment gains, consumption gains and external gains. Holtmann (1973) pointed out in his research that the same educational investment will present greater effect in the older students than the younger ones using South African data in the 1970s. Lanlan Liang (2016) derived the growth model of how educational investments contribute to the economy growth from Solow Model and justified the result statistically using panel data from 1999 to 2013 in 31 areas in China. She argued that, on the one hand, the capital labor investments had a higher payback ratio than material capital investments, while educational investments during elementary education had a higher margin rate than higher education on the other. According to the analysis of the references, personal educational investments during elementary education may bring positive gains on a macro view, while the statistical researches into the micro view, such as their impact on personal studies, rarely exists.

In the minor phase, the monitors of the students, namely their parents, take up the responsibility to choose proper individual educational investments for them. And during the compulsory education period, parents may choose a variety of free-time classes, including physical educational classes (sports, chess, to name a few), art classes (musical instruments, go, calligraphy or painting), English lessons (Public English, New Concept English, Cambridge Young Learners English, etc.), Olympic Maths lessons and Chinese lessons (mainly composition classes). In the evaluation system of compulsory education, the higher or lower competence in P.E. classes or art classes are not rigorously discriminated in the high school entrance examination nor in the university entrance examination, thus these two kinds of classes are not the objects in this research.

Take Beijing as an example. Under the evaluation system of Chinese schools, the subjects of a high school entrance examination includes Chinese, Mathematics, English, Physics and Chemistry. In primary school, pupils have the access to learn the former three. Therefore, it is natural to presuppose that additional educational investments in these three subjects may receive future benefits in not only daily practices, but also in the entrance exam. Especially, the long turn benefits in future exams of early time English courses and Chinese courses learning are widely accepted, yet the act of Olympic Maths learning has been a heated debate ever since the 21st century.

Olympic Maths origins from Soviet Union in the year 1934, and was regarded as the Olympic Games in the field of Mathematics due to its high difficulty. When it comes to the 21st century, the study of Olympic Maths in primary school years is criticized nationally because its extremely high demand in mathematic knowledge and skills. Meanwhile, many an organization claims cooperation with key middle schools when selecting pupils, and parents fear that their

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

child may be dropped behind because of the herding effect, the Olympic Maths classes are never outdated.

As a key factor in primary school individual educational investment, the debates of Olympic Maths have never ceased in the references, most of which are criticisms of all kinds. Beijing Evening News and Legal Evening News (Beijing) quoted the statements of mathematic experts frequently, that there were actually less than 5% of pupils suitable for Olympic Maths learning contra the fact that more than 80% of Beijing pupils were actually studying it. Main stream media presses like China Daily, Southern Weekly and China Education Daily criticized more than once the increasingly-heated Olympic Maths learning ever since the 21st century, and they investigated the fundamental reasons why it came into being through cases or interviews respectively. Qiping Kong (2004) showed worries with the rise of Olympic Maths, and doubted the misleading effect of Olympic Maths on students' logic if they were too young, and argued the courses as contradictions to fundamental rule of study. Dongping Yang (2012), the director of the Chinese Academy of the 21st Century Education expressed his worries to the unnecessary competitions caused by the Olympic Maths studying among pupils. Ailiang Li (2011) pointed out that the plateau or trough of Olympic Maths courses revealed the games between stakeholders, who seriously aggravated the asymmetry of education and thereafter undermined the equality of education. Similar opinion is held by the vice-president of the Education Association in Jilin Province, Wenwu Zou (2013), who criticized Olympic Maths education as a way to contaminate compulsory education's environment, to suppress students' learning interest and to distain teachers' moral. Hao Ge (2017) from Peking University defined what Olympic Maths was from the view of an mathematician, illustrating the underlying reasons of the booming Olympic Maths studying and suggested that students should consider their abilities before making similar decisions of whether to study and who to learn from.

Few people hold positive opinion towards Olympic Maths. Yihao Wang (2016) expressed the worries to irrational Olympic Maths learning, while justified that two groups of students (one has learned Olympic Maths in primary school and the other hasn't) differed in a final-term Maths exam score and Maths confidential level. He argued the possible positive impact of Olympic Maths learning on Maths exam scores.

The reason why parents choose Olympic Maths courses are quite complicated. Firstly, because of the high attendance rate of Olympic Maths, parents may consider themselves to be the targets if not choosing Olympic Maths courses for their kids, thus they tend to do the investment due to peer pressure. Leonardo (2015) argued in a similar research that if students were told to be exposed to public if they would do well at exams, their scores would tend to go downwards to avoid being shown on the bulletin board. It's natural to believe that students are inclined to follow the common routine to avoid the negative impacts when standing out. Secondly, there are numerous parents choosing Olympic Maths for their kids due to their worries of the inequality of education. Although Olympic Maths courses may cost them a great fortune, parents are eager to choose such courses for their kids in expect for the scale effect of knowledge, hoping that their kids can take the advantage of it. Similar references could be witnessed, such as the example of Jamaica in the research of Nigel (2010), where he stated that personal educational investments are least likely to realize equality in education, when the rich is capable of the high expenditure of personal investment that the poor is not. Thirdly, some parents may choose Olympic Maths courses hoping to improve their kids' performance at exams, yet little statistical evidence in favor of this claim could be seen in early references.

According to the references, Olympic Maths investment behavior is greatly influenced by social externalities, and the long term score improvement, which should be the key criteria, is often ignored. When it comes to 2017, the heat of Olympic Maths is never faded, a high money cost and time cost is still paid by parents and their children. Therefore, it is the focus of the entire society to investigate whether the investment may indeed bring score growth in the long run,

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

and if the effect can be presented quantitatively, the research result will not only fill in the blank of empirical analysis of Olympic Maths, but also facilitate the investment decision made by the parents to lessen their anxiety and frustration.

When designing questionnaires for the research (The questionnaire adopted in this research is the same with that used in the former reseatch of Yihao Wang (2016)), Olympic Maths classes are defined as any socially held Maths classes that teach mathematical problem with higher difficulty compared with primary school lessons. Actually, the standard is more loose than authentic Olympic Maths lessons according to Hao Ge (2017), but in order not to cause unnecessary misunderstanding of the testees and represent the real status of social classes, in this research the broader concept is used.

2. Research Methods and Independent Variables

This thesis will take Olympic Maths studying as an example to analyze the profit gained in the future, namely the enhancement of scores, of individual educational investments in primary school. The testees will be drawn from one public middle school. Considering no universal entrance exam is allowed when pupils enter middle school, and university entrance examination is too much beyond the time when investment behavior took place, high school entrance examination will be used as the “underlying asset”.

Questionnaire investigation method is adopted in this research, and other econometrical methods such as OLS and PSM are used before the mathematical model came into being. The testees are drawn from a time-honored middle school in Beijing, whose undergraduate students of grade 2014 and 2015, as well as their parents, are kindly requested to participate in the research voluntarily. The survey finished on the second half of 2012, in which 627 questionnaire pieces are considered effective.

The variables used in this research are based on the references and the results of preliminary experiments, as can be seen in table 1.

Table 1. Names of variables and specific descriptions

Variable Name	Descriptions
GEND	Two-valued variable of gender, male=1, female=0
OLYMA	Two-valued variable of whether the testee has studied Olympic Maths in primary school, yes=1, no=0
COST	Total cost of Olympic Maths studying in primary school, yuan
SEMEST	Total semesters spent on learning Olympic Maths, integer
PAREDU	Sum of parent education level, maximum=12
FAMINC	Family income level, maximum=7
MACONF	Maths confidence level, maximum=5
MAANXI	Maths Anxiety level (reverse scoring), maximum =1
MATFIN	Maths score of district-wide final exam, maximum=120
MATENT	Maths score of high school entrance exam
CHNENT	Chinese score of high school entrance exam
ENGENT	English score of high school entrance exam
PHYENT	Physics score of high school entrance exam
CHEENT	Chemistry score of high school entrance exam
TOTAL	Total score of high school entrance exam

Assumptions are made as follows:

Female students may be more mature during middle school stage, and it is reasonable to believe that female students are more diligent and hardworking than male students. In the pre-going

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

interviews, the class advisor mentioned that female students performed better in terms of daily homework, and were better motivated to ask questions after class. Thus, the coefficient of the variable GEND is expected to be negative, which means female students have higher Maths exam score than male students.

The demand of Olympic Maths courses excess greatly that of primary school Maths classes, and part of its content reaches the difficulty of middle school exams. It is natural to suppose that the Olympic Maths courses may be beneficial to pupils in broadening their horizons and eye-opening. Hence, the coefficient of the variable OLYMA is believed to be positive. Students with Olympic Maths studying experience may have a higher score in future Maths exams.

According to the accumulation effect, it is assumed that the longer time a student spend in Olympic Maths studying, the higher its significance in improving Maths score, i.e. the coefficient of variable SEMEST is positive. Moreover, it is supposed that the more money spent on Olympic Maths, the higher score the students will get, i.e. the coefficient of variable COST is positive.

The education level of one parent is evaluated as follows: primary school=1, middle school=2, high school=3, bachelor's degree=4, master's degree=5 and doctor's degree=6. In practice, the variable PAREDU is the sum of the education level of both parents. It is believed that the coefficient of PAREDU is positive, the higher education background their parents have, the better they are likely to perform at exams.

The family income level is valuated similarly from scale 1 to 7, in which 1 represents a family income per capital less than 1000 yuanes per months and 7 represents an income of more than 10000 yuanes per months. From common sense, children in rather poor family has a stronger will to improve their family's economic status, and possess a stronger motivation to study. It is assumed that the coefficient of FAMINC is negative, which means the better family income level a student has, the poorer he performs.

The Maths anxiety level and Maths confidence level are evaluated as the arithmetic mean of all the answers in the modified standard scales excerpted in references. The anxiety level is marked reversely and the confidence level is numerated positively. The research supposes that the coefficients of both MAANXI and MACONF are positive, which means a lower math anxiety level and higher confidence level will have a positive impact on exam score.

MATFIN is the Maths exam score of all the subjects in a district-wide final exam held in 2012, and was standardized into a full mark of 120 to accompany the high school entrance examination score. Analogous research carried by Wang(2016) found that students who studied Olympic Maths perform better than those who didn't at exams. Thus, in the following research, the model of MATFIN is first inferred to verify the authenticity of the references, and sets an example for the MATENT model.

3. Variable Description and Filtration

The statistical descriptions of the subjects' high school entrance examination scores can be seen in table 2.

Table 2. Statistical descriptions of the high school entrance exam score of the testees

Variables	Obs.	Largest	Smallest	Avg.	Std.	Skewness	Kurtosis
MATENT	447	119	27	96.04	11.84	-2.26	9.63
CHNENT	447	120	54	102.96	8.42	-1.43	4.41
ENGENT	447	119	26	106.04	11.25	-2.99	14.51
PHYENT	447	100	32	82.40	11.31	-0.73	1.08
CHEENT	447	80	20	69.08	7.86	-2.51	9.51
TOTAL	447	518	164	456.51	43.49	-2.39	10.05

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

According to news reports, the average total score of 2014 high school entrance examination in Beijing was 449.57 out of 540 (Data source from: http://zhongxue.hujiang.com/news/zhongkao/p314426/. P.E. score occupies 40 points of the total 580 points, and was deducted in this case), and Maths score reached 91.86 out of 120. The average total score of 2015 high school examination in Beijing was 10 points higher than 2014 and hits 460. The numbers of candidates in 2014 and 2015 in Beijing are almost the same, thus the arithmetic mean would be used as an estimation of the average score level.

According to table 2, all five subjects are left-skewed, and have a significant positive kurtosis, which means there are more students above the average than natural distribution and more students at both ends of the spectrum than normal distribution. It is quite similar to the characteristic of the entirety. Moreover, the average score of this school is more or less the same with the overall, thus the testees could be regarded as a good simulation of the entirety.

The statistical description of the variables can be seen in table 3 and elementary analysis is implemented below.

Table 3. Statistical description of variables

Variables	Obs.	Largest	Smallest	Avg.	Std.	Skewness	Kurtosis
GEND	447	1	0	0.47	0.50	0.14	-1.99
OLYMA	447	1	0	0.85	0.36	-1.91	1.66
COST	446	70000	0	10053	10549.2	2.21	6.58
SEMEST	445	12	0	4.93	3.16	0.13	-0.54
PAREDU	447	12	4	8.05	1.09	0.16	2.66
FAMINC	446	7	2	4.83	1.31	-0.03	-0.66
MACONF	447	5	1	3.62	0.98	-0.60	-0.03
MAANXI	447	5	1	3.65	0.81	-0.65	0.25
MATFIN	438	117	15	95.51	12.15	-1.46	5.68

As for the variable GEND, there are slightly more female testees than male, and is quite close to the entirety. As for the FAMIC, the family income obeys a normal distribution with an average 6000 yuan per month per capita. According to Beijing Municipal Bureau of Statistics, in 2012 the average income per capital in Beijing was 62677 yuan, or 5223 yuan per month. Considering the subject middle school is located in Xicheng District in Beijing, it is reasonable to believe the parents in this school earns more than the average. And in terms of the PAREDU, the average parent education level is above the bachelor's degree.

Elementary analysis concerning the Olympic Maths variable shows that, there are significantly more students with Olympic Maths studying experience than those without, the ratio approaches 5.7:1. Besides, an average 10000 yuan was spent on Olympic Maths studying in an average 5 semesters. The result coincide with the popular belief that most students start their Olympic Maths studying from the 3rd year of primary school. Compared with the income data, nearly 2% of family annual income was spent on Olympic Maths, which is a heavy burden on most families.

The result of Maths anxiety and Maths confidence shows that MAANXI and MACONF basically obeys a normal distribution with a mean of 3.65 and 3.62, respectively. The scale used in the research is reliable.

Because many variables have been proposed, necessary filtration is needed to avoid over explanation. Thus, the dependent variable MATFIN is needed instead of the MATENT in the first step. Those variables that best explains the MATFIN will be considered a priory when putting forward the final model.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

T-test is used to examine whether MATFIN can perfectly represent MATENT, the result of which can be seen in table 4.

Table 4. T-test result of MATFIN and MATENT

Variable	Obs	Mean	Std. Err	Std. Dev	95% Conf. Interval	
MATENT	438	95.8425	0.5651	11.8262	94.7319	96.9531
MATFIN	438	95.5114	0.5808	12.1548	94.3700	96.6529
diff	438	0.3311	0.4542	9.5048	-0.5616	1.2237
	Value	Deg. of Freedom	P>|t|			
T-test	0.7289	437	0.4664			

H_0 supposes that the means of MATFIN and MATENT are the same while the alternative H_1 supposes that their means vary. According to table 4, under the significance of 10%, the null hypothesis cannot be rejected, thus we can say MATFIN is an un-biased estimation of MATENT. The variable best explains MATFIN may be significant explaining MATENT. An OLS regression is made to estimate the MATFIN level, whose result can be seen in table 5.

Table 5. OLS estimation results of MATFIN
(*, ** and *** represents a 10%, 5% and 1% significance respectively. All tables are presented in the same way.)

Independent Variable	Dependent variable	
	MATFIN	MATFIN
GEND	0.977	--
OLYMA	3.229	5.071***
COST	0.000018	--
SEMEST	0.274	--
PAREDU	1.321**	1.245**
FAMINC	-0.239	--
MACONF	2.782***	2.842***
MAANXI	1.502	1.460*
cons	65.908***	65.721***
observations	436	436
R-square	0.140	0.127

In the correlation of the second column, only PAREDU and MACONF are significant variables and is contrary to the common sense. Thus modifications are implemented by erasing insignificant variables, namely COST, FAMIC, GEND and SEMEST. The new correlation result can be seen in the third column and the rest four variables are all significant. Considering the R-square only drops slightly, the model in the third column is more reasonable. The model of MATFIN can be written below:

$$\underset{(4.730)}{\mathrm{MATFIN} = 65.72} + \underset{(0.753)}{1.46\mathrm{MAANXI}} + \underset{(0.500)}{1.24\mathrm{PAREDU}} + \underset{(0.623)}{2.84\mathrm{MACONF}} + \underset{(1.503)}{5.07\mathrm{OLYMA}}$$

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

4. Maths High School Entrance Examination Score Model And Basic Tests

4.1. Model construction

In the above chapter, four significant variables have been found in the MATFIN regression. In this chapter, the importance of the above-mentioned variables will be tested primarily.

Individual regression is carried on between MATENT and all independent variables, as can be seen in table 6.

Table 6. Result of single-factor regression of MATENT

Variable	t	p-value	R-square
GEND	0.67	0.504	0.001
OLYMA	3.28	0.001**	0.046
COST	3.58	0.000**	0.024
SEMEST	3.21	0.001**	0.028
PAREDU	2.85	0.005**	0.026
FAMINC	-0.31	0.76	0.000
MACONF	5.54	0.000**	0.061
MAANXI	4.07	0.000**	0.067

The regression result shows that GEND and FAMINC are not significant factors in evaluating Maths high school entrance examination score, and future model will no longer encompass these two factors. A comprehensive regression model is shown in table 7.

Table 7. OLS estimation results of MATENT

Independent Variable	Dependent variable
	MATENT
OLYMA	4.412*
COST	0.000049
SEMEST	0.136
PAREDU	1.369**
MACONF	1.767**
MAANXI	2.317*
_cons	65.293***
Obs.	445
R-square	0.138

From table 7, under 5% significance, PAREDU and MACONF are significant and under 10% significance, OLYMA and MAANXI are also significant. The other two variables, namely COST and SEMEST, are no longer significant and are deducted from the final model. The new regression result can be seen in Table 8.

Hereby all four variables in table 8 are significant under 5% significance, and R-square is ideal in the new model. Thus the model of Maths high school entrance examination score can be written as follows:

$$\mathrm{MATENT} = 64.847 + 2.37\mathrm{MAANXI} + 1.36\mathrm{PAREDU} + 1.87\mathrm{MACONF} + 5.71\mathrm{OLYMA}$$
$$(6.146) \quad (1.199) \quad (0.542) \quad (0.731) \quad (2.015)$$

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

Table 8. Final OLS estimation results of MATENT

Independent Variable	Dependent variable
	MATENT
OLYMA	5.713***
PAREDU	1.357**
MACONF	1.869**
MAANXI	2.371**
_cons	64.848***
Obs.	447
R-square	0.138

4.2. Multicollinearity test

Multicollinearity means the independent variables are highly related to each other, which will cause an accuracy bias in the model. To test its existence, multicollinearity test is carried out with result shown in table 9.

Table 9. Multicollinearity test result

Variable	VIF	1/VIF
MAANXI	1.28	0.779
MACONF	1.28	0.783
OLYMA	1.02	0.979
PAREDU	1.02	0.984
Mean VIF	1.15	

No variance inflation factor is greater than 10, thus there are no multicollinearity in this model.

4.3. Endogeneity tests

Endogeneity means the independent variables are correlated with the residual, and fails to meet rigorous exogeneity. The result of endogeneity test can be seen in table 10.

Table 10. Endogeneity test result

	Deg. of Freedom 1	Deg. of Freedom 2	F-Value	P>|t|
F-test	3	439	1.11	0.3455

F-statistics is 1.11, and a p-value of 0.3455 means under 10% of significance, the null hypothesis that there are no omitted variables cannot be rejected. Thus, it proves that there are no higher degree of independent variables omitted in the model.

5. Propensity Score Matching (PSM) —A Counterfactual Estimation

5.1. Signification of PSM and counterfactual estimation

In chapter 4, a regression model to describe MATENT with independent variables OLYMA, MAANXI, MACONF and PAREDU is proposed. Although there are no multicollinearity problems or higher degree endogeneity problems, the self-selection bias, however, is a problem left unsolved. Self-selection bias describes the endogeneity caused by the dependency between the variables. The model will no longer be un-biased with self-selection bias, and the result will no longer be statistically significant.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

Take the MATENT model as an example. Although whether a student learns Olympic Maths may influence his exam score, it is highly possible that he chose to study Olympic Maths when he was young because of his lower level of Maths anxiety, higher level of Maths confidence or higher exam grade. The reliability of the model will be seriously endangered before the self-selection bias is eliminated. Mads Meier Jæger (2009) illustrated that the cultural capital level, along with the ability to transfer and the ability to receive influence the score of students after testing high school students from Denmark. In the research, parents with higher education level have more cultural capital than the others, and whoever has higher cultural capital transfer ability is more likely to choose Olympic Maths for their kids. The above-mentioned influence need to be eliminated by the experiment or by other methods.

One effective way to solve the self-selection bias is to observe the difference in one subject when studying and not studying Olympic Maths simultaneously, thus all other factors will be held equal. However, it is impossible that a student exist two conditions "Has studied Olympic Maths" and "Hasn't studied Olympic Maths" at the same time, making the difference of the two conditions unable to be observed. If two groups of students are used to compare their score, other factors are hard to be kept strictly equal. Hence, a counterfactual estimation is implemented in solving the case.

The methodology of counterfactual estimation is described as follows: The composition and number of the subjects remain the same. The subjects are divided into two groups by the aimed two-valued variable OLYMA, whose propensity to choose "yes" or "no" is explained by the rest of the independent variables. After the explanation, the experimental group and the observation group are matched accordingly, by ratios of 1:1, 1:2 or 1:3, and was paired into different levels (stratification). In each level, the only difference between the subjects is the difference in aimed variable, then the difference in dependent variable can be tested. Later, average treatment effect (ATT) can be valued through a weighting mechanism.

In this research, ATT embodies the "counterfactual mechanism". The reason why subjects in the observation group have never studied Olympic Maths varies, yet a stratification method could eliminate all other potential differences in subjects between the observation and the experimental group, after which the pure influence of studying Olympic Maths on observation group subjects will be accessible.

How to implement the stratification method refers to Propensity Score Matching (PSM) method. PSM was put forward in 1983 by Rosembaum and Rubin, is "an algorithm to match the experimental group and the observation group based on the conditional probability and propensity score under certain observable conditions". After the computation, the propensity score of the subjects in one strata are the same or in the approximate interval, and they can be regarded as the same except the aimed independent variable. Thus the problem of endogeneity is solved.

Robin Wright (2008) proved in his research that PSM is an effective simulation of randomized sampling, and it is the best way to erase error when the samples or funds of the research is limited. Therefore, PSM method will be used to test the model presented above.

5.2. Principle of PSM

Propensity score is calculated by the following formula:

$$\mathrm{P}(X_i) = \Pr\{exp_i = 1|X_i\}$$

In which X is matching variable, and the estimation of $\mathrm{P}(X_i)$ is the possibility of the subject's investment on Olympic Maths courses.

In practice, linear model Probit or Logit is often used to simulate (Linear model Probit or Logit is defined as follows: $\mathrm{P}(\mathrm{Y}=1)=\mathrm{f}(\mathrm{X})$, where Y is a two-value variable representing the occurrence or nonoccurrence of the event. The model indicates that the possibility of

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

occurrence is a function of a dependent variable that obeys normal distribution. Especially, if f(X) is an accumulated distribution function F(X), the model is Logit. There are no major differences between the two and they can be used interchangeably). In the study, probability estimation P_i of subject i from experimental group and estimation P_j of subject j from observation group are generated using model Probit. The average treatment (ATT) effect of Olympic Maths learning on Maths high school entrance exam score is calculated as follows:

$$\mathrm{ATT} = \frac{1}{N} \sum_{i \in (exp=1)} \left[MATENT_i - \sum_{i \in (exp=0)} W_{ij} MATENT_j \right]$$

In the formula, N equals to the number of subjects in the study, W_{ij} represent the weight given to subject j when his exam score $MATENT_j$ is replaced by subject i's exam score $MATENT_i$.

When the variable means of the subjects in two groups are the same, while all other variables that affect the exam score are controlled after matching, the PSM can attain an objective ATT. When choosing different matching methods, the weight functions W_{ij} are different. In the thesis, Kernel matching method is adopted consulting other literatures, whose function is written below:

$$W_{ij} = \mathrm{K}\left(\frac{P_i - P_j}{H}\right) \div \sum_{j \in (exp=0)} \mathrm{K}\left(\frac{P_i - P_j}{H}\right)$$

In which $\mathrm{K}\left(\frac{P_i - P_j}{H}\right)$ obeys Gaussian normal distribution, and parameter H is bandwidth parameter.

As for the matching variables, the choice of X are caused by all the rest independent variables, i.e. PAREDU, MAANXI and MACONF.

5.3. Realization of counterfactual matching method and results

According to principles of PSM, Probit function is used when estimating the probability of studying Olympic Maths. Due to the limited subjects' number of observation group, the comparison ratio is set at 1:1 and the maximum gap in the same strata is set at 0.05 between different propensity scores.

The t-test result of the Maths exam score after matching is shown in table 11. The balance test of the rest three variables are shown in table 12.

Table 11. Treatment effect of PSM

Dependent variable	Treatment Effect	Treatment Group	Control Group	Gap	Standard Error	**T value**
MATENT	Unmatched	97.1216	90.1159	7.0057	1.5157	**4.62**
	ATT	97.0664	89.8803	7.1861	2.7926	**2.57**

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

Table 12. Result of balance test of the rest variables

Variable	Unmatched Matched	Mean		T-test	
		Treated	Control	t-value	P>|t|
MAANXI	U	3.6846	3.4852	1.88	0.061
	M	3.6831	3.6845	-0.02	0.982
MACONF	U	3.6566	3.4696	1.45	0.147
	M	3.6521	3.6284	0.36	0.721
PAREDU	U	8.1005	7.7536	2.44	0.015
	M	8.0798	8.0878	-0.11	0.913

Due to the limited number of the observation group, a 1:1 matching method caused the loss of many subjects in the experimental group. As can be seen in table 12, after the matching, 69 subjects in either groups have the same level of Maths anxiety, Maths confidence and parent education level, the distinction amongst them has been erased after the matching phase, along with the disappearance of the endogeneity.

As can be seen in table 11, before the disposure an apparent distinction could be seen between two groups of subjects, and after the matching the distinction remains significant on a 1% significance. Then it is appropriate to say there are endogeneity between OLYMA and the rest three variables, yet the endogeneity hardly influence the significant impact made by learning Olympic Maths on Maths exam score.

In a word, the Maths score model presented in chapter 4 is credible statistically.

6. Discussion of Maths High School Entrance Exam Score Model

6.1. Realistic significance

Recall the Maths high school entrance exam score model presented in chapter 4.

$$\mathrm{MATENT} = 64.847 + 2.37\mathrm{MAANXI} + 1.36\mathrm{PAREDU} + 1.87\mathrm{MACONF} + 5.71\mathrm{OLYMA}$$

The PSM in chapter 6 has proved its authenticity and it is appropriate to discuss its realistic significance.

According to the model, when the sum of parent education level increases by one unit, the Maths score will go up by 1.36 points, coincide with the former assumption. The Maths subject is difficult for its logical thoughts, and loads of exercises are required to consolidate whatever learned at school in the spare time. Though it is difficult for middle school students to be aware of it, the instruction and supervision from parents are vital to help them realize it and learn better. The higher education level their parents are, the more likely students are instructed systematically and scientifically.

When the Maths anxiety level score increases by one points (By reverse scoring, it equals to the decrease of real anxiety level), their Maths high school entrance exam score will increase 2.37 points, as expected in the previous assumption. The Maths anxiety scale evaluates the degree of anxiety when facing Maths exams or doing Maths exercises, whose effect is intensified in the test-oriented education system. The fear of difficulties before and after the exams seriously undermines their true horizon. A decrease in anxiety helps to improve the Maths exam score.

When the Maths confidence level score increases by one point, the exam score will increase 1.87 points, as expected before. The Maths confidence scale evaluates the self-recognition when studying Maths. Those with higher confidence level tends to spend more time solving

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

mathematical problems, and they will spare more effort trying the challenging problems. Along with higher self-expectations, they can accumulate more as days go on compared with their peers, and are more likely to solve out the problems with high degree of discrimination in the recruitment exams.

The most valuable information received from the model is, those who studied Olympic Maths in primary school outperform the others by 5.71 points, ceteris paribus. The model confirms the long-run profit of Olympic Maths investment, and embodies the profit in high school entrance exam score quantitatively.

It is noteworthy that the cost of Olympic Maths studying and the semester of studying are not significant statistically although their coefficient is positive. In the experimental group, the economy of scale of cost and time involved in studying Olympic Maths is not supported by the statistics. This seems contrary to the common sense, but on the other hand, the result may be biased when the subjects reported it in the questionnaire, as it had been years after they ceased studying Olympic Maths and the exact money paid may be forgotten. As for the period of studying, the influence of Olympic Maths may vary among those students who has studied the same semesters, since the time when they started, the time when they stopped and whether they studied it consecutively are all influential factors to the effect. Since none of them is considered in the regression, the accuracy of the variable SEMEST is undermined. Thus, further investigation are needed to work out the authentic effect of the cost and time of studying Olympic Maths.

Contrary to the assumptions made, however, the gender and family income level show no statistically significant impacts on Maths high school entrance exam score. Although female students may study harder than boys, but male students may compensate their comparatively poor dedication to study with other advantages. Similarly, though students from relatively poor families have stronger ambition, it is likely those from wealthy families have better access facilitating their study, like hiring family tutors. Thus, the gender and family income differentiations are hardly reasons to explain the final score.

6.2. Model defects

Although the model explains the Maths high school entrance exam score, there are flaws in the model, both systematic and unsystematic.

Systematic defects are flaws that cannot be avoided by the increase in the subjects' number. Firstly, the ideas conveying bias matters. This survey concerning family information, such as the parent education background, family income level and cost of Olympic Maths studying, has something to do with privacy. In order to match the survey result of the two questionnaires and the high school entrance examination score, the questionnaires adopted in the research are name-required questionnaires. We cannot rule out the possibility that the subjects provide inauthentic answers in fear for information leak. Secondly, the missing variable bias. The observations explained in the model occupy around 15% of total subjects, and there are significantly important factors, e.g. the teaching level of the head teacher, the harmony of student's family or even the distance from home to school, missing in the model. The investigation of these factors is beyond the accessibility of the research. Due to its origination, the systematic risks are considered acceptable in the research.

Unsystematic risks are risks that can be ruled out by increasing the numbers of testees, in this research it refer to the sample error. The subjects in this study are from two consecutive grades in one middle school, although their exam score distribution and the real distribution of all candidates are alike, they can hardly represent the real cases of all two hundred thousand middle school students in Beijing on a macro view. For those poor middle schools, there are less students studied Olympic Maths, while in those top middle schools, such as Middle School Affiliated to Renmin University or Beijing NO.4 middle school, students had been filtered by

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

their Olympic Maths studying experience when they entered, and almost 100% of them outperform their peers from other middle schools in Olympic Maths, making the Maths exam score model no longer effective in such schools. The authenticity can be improved by choosing more subjects in a variety of schools to erase the sample error. Though such method is unavailable in this research, future investigations or similar researches may avoid unsystematic problems.

7. Influence of Olympic Maths Investment on Other Courses In High School Entrance Examination

7.1. Explanation of regression result

After the discussion in Chapter 6, this chapter continues to discuss the influence of Olympic Maths studying. The impact, if had any, on the rest four courses in the high school entrance examination of Olympic Maths studying is also valuable.

In this chapter, variables MAANXI and MACONF are no longer necessary in our discussion as it specifies the confidence and anxiety in Maths. PAREDU, on the other hand, is still valuable and will be incorporated. Specifically, the effect of gender is an interesting point to study, thus variable GEND is also included. The results of the regression are presented in table 13.

Table 13. OLS regression results of other courses in high school entrance exam

Independent variable	Dependent Variable				
	CHNENT	ENGENT	PHYENT	CHEENT	TOTAL
OLYMA	2.791**	3.399**	3.336*	3.279**	19.272***
GEND	-3.115***	-3.983***	-0.538	0.248	-6.780*
PAREDU	0.648	1.765***	1.333**	0.761*	5.993***
R-square	0.056	0.075	0.031	0.038	0.059

The characteristics of the regression are summarized as follows:

1. Olympic Maths studying is beneficial to all four courses and the improvement is statistically significant. The effect differs among subjects. The total effect is a 19.272 points increase compared with those who didn't study Olympic Maths.

2. Gender difference is not a significant factor in science subjects, but is apparently significant in art subjects. Female students perform significantly better than male students in Chinese and English, but no apparent difference could be witnessed in Maths, Physics or Chemistry. As for the total score, female students perform 6.78 points higher than male students, and the discrepancy is marginally significant under 10% of significance.

3. Parent education level leaves a significant impact on all courses except Chinese. As for the total score, 1 unit's increase in PAREDU (for example, one of the parents gets a master's degree rather than bachelor's degree) will bring a 5.993 points increase in the total score.

Although the influence discussed above are significant statistically, the R-square value is relatively low, the majority of the observations are left unexplained. As future research is required, no exam score models of other courses will be presented in this thesis.

7.2. Realistic significance

The regression result in table 13 gives a clear presentation on the influence of OLYMA, GEND and PAREDU. The explanation to these variables are realistically significant.

7.2.1. Parent education level

Parent education level leaves a significant impact on all subjects, include Maths subject as have been discussed in early chapters. Those parents who has higher education background knows

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

the importance of learning better than other parents, and a high education level represent their comprehensive culture level, making them easier to cultivate their children's good learning manner and study habit. The regression points out that the influence is significant not only on courses with logical demand, but also on courses that demands mass accumulation (English), requires high understanding of experiment operations (Physics) and needs mature basic training (Chemistry).

Statistically, parent education level influence most on English (1.765 points), followed by Maths (1.367 points) and Physics (1.333 points). The impact on Chinese is not significant. The research suggest that Chinese requires more comprehensive skills, include academic reading, narrative writing, ancient prose understanding, etc. Typical skills and proper professional guidance is the base of high score, and somehow diluted the importance of parents' unprofessional cultivation. English, however, can be easily improved with sufficient vocabulary bank, and the parent instructions are especially vital to their kids' accumulation.

7.2.2. Student gender difference

The difference in gender partly decides the way of thinking and study attitude, whose effect is broadened in middle school. The exams in primary school are tests for completion, while the exams in middle school are tests for selection. Students may only get a higher score with better study attitude and sufficient revision time. Compared with male students, female students may be more diligent as have been suggested in early chapters, thus they are easier to get a higher score in Chinese and English. As for the science subjects, the weakness in diligence may be compensated by their advantage in logical thinking, thus the boys can perform equally with girls in Maths, Physics and Chemistry.

7.2.3. Experience of Olympic Maths studying

The result shows that students with Olympic Maths studying experience outperform their peers in all five subjects. Since the content of Olympic Maths courses has nothing to do with the rest four subjects, the relationship is somehow doubtful. However, though it is difficult to explain the result through the knowledge itself, the underlying mechanism is more complicated.

As has been mentioned, the exams in middle schools are selection-oriented, the pattern of exam is quite similar to that of Olympic Maths courses. The knowledge required in Olympic Maths courses is demanding, and their exams are treated as a criteria when top middle schools select students. Due to the same function of the exams, students will study as hard as they can when preparing for Olympic Maths exams, and they are more skillful and familiar with the selection-oriented exams after entering middle schools. Therefore, they're more experienced in arranging their studies and balancing their leisure time than those who didn't prepared for such exams in primary school. Such first-mover advantage will benefit all their studies.

8. Suggestions on personal Olympic Maths investment

As have been mentioned at the beginning of the research, when people make educational investment decisions, they will analyze the future outcome and have it compared with the cost of money and time according to personal preferences. If the profit turned out to be larger than the cost, a rational person will certainly exercise the investment.

Considering that all the testees in the research are juveniles, their educational investment decisions are made by their parents. The cost of Olympic Maths studying, from their parents' point of view, includes the time spent on the way to school and home every weekend, the registration fee of each course and the occupation of leisure activities. The cost of the students includes not only the time spent, but also the utility loss if the same period of time is used for playing or other hobbies. The profit of such an investment, includes not only preferential

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

policies when entering middle school, but also a long run reward at high school entrance examination.

The coverage ratio of Olympic Maths in this research is over 80%, and it is reasonable to believe a similar ratio could be witnessed city-wide. The beneficiaries of preferential policies are quite limited, and the majority of the students may receive no short-term benefit after studying Olympic Math. Their parents expressed a similar worry in the questionnaire, as they claim their major wish of choosing Olympic Maths investment is to get their child an admission ticket to the key middle schools. Although no statistical investigation was made analyzing the long term reward in high school entrance exam scores, this study provides a good reference for future parents. The research confirms the long-run benefit of Olympic Maths learning in primary school, giving future parents more confidence in making such decisions for their kids.

In conclusion of the model result and relevant references, the research suggests that if the child is not inimical with Olympic Maths courses, and if he can spare more time after accomplishing his daily routine, it is recommendable for parents to arrange Olympic Maths courses moderately. From such courses students may broaden their horizons and enhance their independence to study, and will be more prepared for the future exam-oriented education.

It is noteworthy that although Olympic Maths studying is contributable to future high school entrance exam score, the impact didn't present a return-to-scale effect along with the increase of money involved, thus parents should be careful not to fall into "over-investment trap". Meanwhile, the opportunity cost of Olympic Maths studying involves the pleasure of taking other art courses, a strong will and vigorous body brought by physical exercise. A balanced development is needed when concerning Olympic Maths studying. Moreover, Chinese and English subjects are all time consuming, and in primary school time enough attention should also be paid to these courses, to prevent an asymmetry in all the courses.

References

[1] Cheung, A.K.-L. (2014) "Hiring domestic help and family well-being in Hong Kong: A propensity score matching analysis," Journal of Comparative Family Studies, 45(4), 475–495. https://doi.org/10.3138/jcfs.45.4.475.

[2] Holtmann, A.G. (1973) "The timing of investments in human capital: A case in education," Public Finance Quarterly, 1(3), 300–306. https://doi.org/10.1177/109114217300100303.

[3] Li, J. (2004) "'I learn and I grow big'': Chinese preschoolers' purposes for learning," International Journal of Behavioral Development, 28(2), 116–128. https://doi.org/10.1080/01650250344000325.

[4] John, L. et al. (2007) "The use of propensity scores as a matching strategy," Research on Social Work Practice, 18(1), 20–26.https://doi.org/10.1177/1049731507303958.

[5] Bursztyn, L. and Jensen, R. (2015) "How does peer pressure affect educational investments?," The Quarterly Journal of Economics, 130(3), 1329–1367. https://doi.org/10.1093/qje/qjv021.

[6] Jaeger, M.M. (2009) "Equal access but unequal outcomes: Cultural Capital and educational choice in a meritocratic society," Social Forces, 87(4), 1943–1971. https://doi.org/10.1353/sof.0.0192.

[7] Brissett (2010) "Jamaica's policy discourse in the age of globalization: Framing education as (private) investment," The Global South, 4(2), 197-217. https://doi.org/10.2979/globalsouth.4.2.197.

[8] Kuhn, R., Everett, B. and Silvey, R. (2011) "The effects of children's migration on elderly Kin's Health: A counterfactual approach," Demography, 48(1), 183–209. https://doi.org/10.1007/s13524-010-0002-3.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202307_6(7).0016

[9] ROSENBAUM, P.A.U.L.R. and RUBIN, D.O.N.A.L.D.B. (1983) "The central role of the propensity score in observational studies for causal effects," Biometrika, 70(1), 41–55. https://doi.org/10.1093/biomet/70.1.41.

[10] Fan, X. and Nowell, D.L. (2010) "Using propensity score matching in educational research," Gifted Child Quarterly, 55(1), 74–79. https://doi.org/10.1177/0016986210390635.