Educational Research and Human Development
ISSN: 3007-6943
DOI: https://doi.org/10.61784/erhd3016

AN ANALYSIS OF FACTORS AFFECTING STUDENTS' PERFORMANCE IN MATHEMATICS IN A SECONDARY SCHOOL BASED ON QUANTILE REGRESSION

XiaoXiao Wu, SiMei Pan*, Chia-Yi Luo, Lin Yun, TsungXian Lin, WenChao Pan
School of Management, Guangzhou Huashang College, Guangzhou 511300, Guangdong, China.
Corresponding author: SiMei Pan, Email: 324827824@qq.com; pansm@gdhsc.edu.cn

Abstract: The study of factors affecting secondary school students' performance is an important topic for educators, especially at a time when the state pays special attention to education reform. This paper uses quantile regression to explore in detail how the relevant indicators affecting secondary school students' performance affect students' performance and the degree and direction of influence at each quartile, in order to provide relevant educators as a reference for decision-making on education reform. The results of the study show that weekly study time is significantly underestimated by OLS in the lower quartiles, which means that only about 2 hours of study time per week can improve academic performance; extracurricular activities are also significantly underestimated by OLS in the higher quartiles, which means that increasing the number of extracurricular activities can improve academic performance.
Keywords: Math achievement; Influencing factors; Quantile regression

1 INTRODUCTION

Secondary school students' performance in mathematics is very important. Mathematics is not only a key subject to bridge the gap at the secondary school level, but also directly affects students' overall performance and future development. First of all, math plays a crucial role in middle school. It is not only the foundation of natural science, but also exercises students' logical thinking and problem-solving skills. With the gradual deepening of learning content, from basic algebra and geometry to complex equations and functions, math requires students to have an active mind and the ability to continue learning. Good or bad performance in mathematics often intuitively reflects the ability of students in continuous learning. Students with strong thinking ability can not only achieve excellent results in mathematics, but also transfer this ability to other subjects. Second, the gap in math scores is especially pronounced in the midterm exams. Good or bad math scores have a direct impact on whether students can enter a good high school, because math scores can easily widen the score gap. Excellent math scores can lead to high scores in the midterm, while students who do not do well may fail. Therefore, good or bad math scores are crucial for performance in the secondary school exam. This paper adopts the data related to students' mathematics course performance in a secondary school to study and understand which important indicators affect the semester grades of the students in this school, in order to provide relevant educators in China as a reference for teaching reform.

In the literature related to the study of secondary school students' performance in mathematics，Luo Qiang and Yu Feifei conducted a study on the academic quality monitoring data of junior high school students in S city, and the results of the study showed that boys had a more obvious academic advantage in difficult topics; the math scores of girls in the low subgroups were significantly higher than those of boys, and the math scores of girls in the middle and high subgroups were significantly lower than those of boys[1]. Yue Qiming studied the core journals with the theme of "education informatization" in China since 2018, using China Knowledge Network (CNKI) as the data source. The analysis shows that the reform of teaching mode, information technology and big data to promote the development of school education to personalization and intelligence is the focus of the current research on education informatization; accelerating the construction of information infrastructure projects, optimizing high-quality digital educational resources, and the in-depth fusion of information technology and education and teaching are the key directions of China's education informatization[2]. Fan Shuyuan and Xiong Yonghong conducted a study on the characteristics of the teaching mode of experimental courses, and the results of the study showed that due to the application of teaching forums, a learning group was formed in the virtual environment with the experimental project as the topic, which facilitated the communication between teachers and students, cultivated the students' interest in learning, and also brought teachers and students closer to each other and to each other[3]. Li Wan analyzed the needs and problems of mathematical modeling teaching based on the concept of "1+X" and showed that it is recommended that mathematical modeling teaching in colleges and universities should be closely linked with professional courses, introduce real cases, strengthen mathematical modeling competitions and practice sessions; adopt flipped classroom and blended teaching, strengthen group cooperation and project-based teaching; and adopt the concept of "1+X" and the concept of "1+X" in mathematical modeling teaching. The use of flipped classroom and blended teaching, strengthening group cooperation and project-based teaching; focusing on the combination of process evaluation and result evaluation, and paying attention to the evaluation of students' ability and innovation quality[4]. Deng Wenxin et al. conducted a study on the research characteristics of China's vocational education informatization in the past 10 years, which showed that under the new situation of the development of the times, digitalization and intelligence of education may become a new

research hotspot; the development of the research needs to strengthen the cooperation of researchers with each other, and to continuously innovate the design of the research and research methodology[5]. Liu Yong conducted a research on the educational teaching evaluation system based on artificial intelligence technology through the development and application of artificial intelligence technology, and the research shows that it provides relevant content references for the innovative application and practical exploration of artificial intelligence technology in higher education teaching evaluation[6]. Bie Dunrong conducted a study on the new major variables that promote the development of university education and teaching, and the study shows that in the face of the opportunities and challenges arising from the application of AI technology in university education and teaching, teachers should actively learn and master AI education technology to develop education and teaching methods that cultivate students' higher-order abilities and qualities; students should actively learn AI education technology to obtain targeted services and assistance and realize their own personalized learning. Students should actively learn AI education technology to get targeted services and help to realize their own personalized learning[7]. Kang Yueyuan, Zhang Junhong and Song Chunli studied the hotspots and future development trends of international mathematics education research in the context of the era of digital intelligence, which showed that evidence-based research in mathematics education focuses on evidence-based research in mathematics education, wisdom focuses on the development of intelligent teaching and learning ecosystems, ubiquity demonstrates the extensive use of digital intelligence in mathematics education, and ethics emphasizes the ethical principles that must be taken into account in the application of digital intelligence technologies. Ethical principles that must be emphasized[8]. Li Tieying and Wang Yunfeng conducted a study on the process of promoting the integration of Civics and Political Science courses in universities, secondary schools and elementary school, which showed that there are problems in the construction of teaching resource libraries, and efforts should be made to promote the integration of teaching resources in Civics and Political Science courses in universities, secondary schools and elementary school in the aspects of strengthening the top-level design, developing high-quality teaching resources, strengthening the efforts of joint construction and sharing, and creating a group of excellent construction teams[9]. Song Qijie for the study to tourism secondary vocational education teaching reform as a starting point, from the curriculum construction, talent supply, teacher construction, teaching resources construction four aspects of the study, the study shows that the digital transformation background of digital transformation of secondary vocational school tourism professional digital transformation path; so that secondary vocational education and the new period of social development synchronization, to help China's new era of talent supply system[10].

In summary, domestic scholars for education students domain knowledge this paper is based on a specific environmental context as a comprehensive indicator of data analysis, while the evaluation results to do input redundancy and output insufficiency analysis, for the development of theoretical and practical basis.

2 RESEARCH METHODOLOGY AND SELECTION OF INDICATORS

2.1 Quantile Regression Model

Quantile regression (English: Quantile regression) is a modeling method for estimating the linear relationship between a set of regression variables, X, and the quantiles of the explanatory variables, Y. It is one of the methods of regression analysis, which was originally proposed by Koenker Roger and Bassett Gilber5t Jr in 1978.

The survey included: basic information (gender, age, parents' education level, etc.), time from home to school, weekly study time, presence or absence of extracurricular activities, desire for higher education, quality of family relationships, free time after school, number of absences, and semester grades. The higher the value of time from home to school, weekly study time, quality of family relationships, free time after school, and semester grades, the higher the level, the presence of extracurricular activities is assigned according to the length of time, and the desire for higher education is assigned according to the student's personal wishes.

In this paper, we take three tertiles of q25, q50, and q75 to perform quantile regression analysis for independent variable X, and dependent variable Y. $P < 0.05$ was taken as significant.

2.2 Indicator System and Data on Indicators

Table 1 Indicators Related to the Mathematics Curriculum of Students in a Secondary School

variant	assign a value to something
sex - sex of the student	F=Female, M=Male
age - age of the student	15-22 years
Medu-Mother's Education	0=None, 1=Primary education (grade 4), 2=Grades 5-9, 3=Secondary education, 4=Higher education
Fedu-Father's Education	0=None, 1=Primary education (grade 4), 2=Grades 5-9, 3=Secondary education, 4=Higher education
traveltime - time from home to school	1 = <15 minutes, 2 = 15-30 minutes, 3 = 30 minutes - 1 hour, 4 = >1 hour
studytime - weekly study time	1= <2 hours, 2=2-5 hours, 3=5-10 hours, 4=>10 hours
activities - extracurricular activities	Yes or No
Higher - want to receive higher education	Yes or No

Famrel - quality of family relations	1=very poor, 2=poor, 3=fair, 4=good, 5=very good
freetime-after school free time	1=very little, 2=somewhat little, 3=no more, no less, 4=somewhat more, 5=very much
Absences -- number of absences	From 0-93
G2 - Semester grades	From 0-20

3 Empirical Research

3.1 Normal Distribution Tests for Explained Variables

In this paper, the first test of normal distribution of the explanatory variables (Y) is conducted by using two commands in stata (sktest and swilk), for the explanatory variables. The hypothesis of the test is that H0 is that the explanatory variables conform to normal distribution; H1 is that the explanatory variables do not conform to normal distribution. A test result of less than 0.05 means that H0 is rejected, which means that the explanatory variables do not conform to normal distribution. From the analysis results, it is found that the swilk test result is 0.00001, which is significantly less than 0.05, so the explanatory variables do not conform to normal distribution. Then from the test results of sktest observed p-value of 0.0003, also significantly less than 0.05, the explanatory variables also do not conform to the normal distribution, so this paper performs quantile regression analysis to estimate this non-normally distributed sample data.

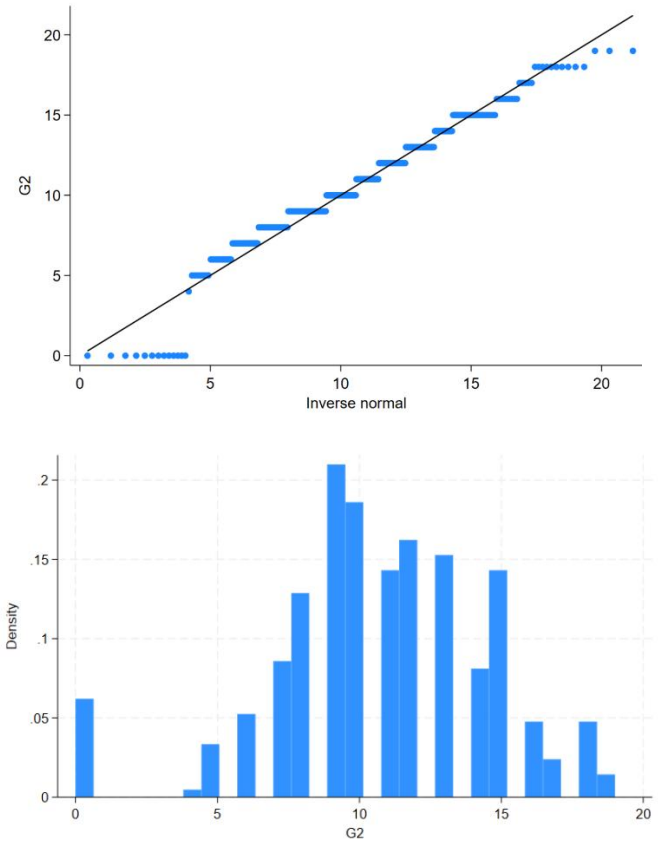

Figure 1 Graphical Test of Normal Distribution of Explanatory Variables

And then by the normal distribution of the explanatory variables graphical observation, Figure 1 in the upper graph of the blue points closer to the diagonal represents the sample closer to the normal distribution, as seen by the blue points in the figure in the left and right sides of the points are not on the diagonal, so it is clear that the variable final grades are clearly not normally distributed. In addition, the histogram data in the lower panel of Figure 1 is also clearly skewed to the right, and there are many samples that are skewed to the far left, which is clearly not a normal distribution pattern, so once again, it can be demonstrated that the final grades of the variable clearly do not follow a normal distribution, so we must perform quantile regression analysis to explore the characteristics of these outliers.

3.2 Quantile Regression Analysis

Table 2 Interquartile Coefficient Test for Final Grades

mould	OLS			0.25 quartile			0.50 quartile			0.75 quartile		
variant	ratio	T-value	statistically	ratio	T-value	statistically	ratio	T-value	statistically	ratio	T-value	statistically

			significant			significant			significant			significant
sex	-1.02	-2.31	**	-0.96	-1.6	-	-1.20	-2.07	**	-0.72	-1.27	-
age	-0.16	-0.90	-	-0.00	-0.01	-	-0.18	-0.96	-	-0.23	-0.80	-
medu	0.39	1.57	-	0.12	0.38	-	0.17	0.47	-	0.68	2.00	**
fedu	0.04	0.16	-	0.30	1.02	-	0.09	0.23	-	-0.00	-0.01	-
traveltime	-0.69	-2.21	**	-0.16	-0.33	-	-0.64	-1.92	*	-0.64	-1.34	-
studytime	0.46	1.79	*	0.59	1.65	**	0.22	0.72	-	0.46	1.24	-
activities	-0.01	-0.02	-	-0.31	-0.59	-	0.17	0.33	-	0.09	0.18	-
higher	2.84	-2.86	***	1.82	-1.01	-	2.43	-2.02	**	2.72	-2.2	**
internet	-0.24	-0.40	-	0.04	0.04	-	-0.25	-0.37	-	-0.30	-0.39	-
famrel	-0.08	-0.35	-	0.28	0.94	-	-0.15	-0.49	-	0.02	0.09	-
freetime	-0.10	-0.47	-	-0.25	-0.87	-	-0.06	-0.21	-	-0.06	-0.21	-
absences	0.01	0.19	-	0.00	0.07	-	0.00	0.02	-	-0.03	-0.96	-
Pseudo R^2	0.0711			0.0854			0.0664			0.0710		

Note: One * indicates a p-value of less than 0.1; two * indicates a p-value of less than 0.05; and three * indicates a p-value of less than 0.01.

Gender (sex) was found to be significant with a p-value of less than 0.05 when analyzed by os, but it was not clear at which quartile it was significant, but from the quantile results we know that it was significant at the 0.5 quartile. From Figure 2 and Table 2, it is found that the more skewed females are, the higher the scores are, probably because females generally mature earlier and are more able to be still to accomplish their goals, whereas males may be more active and are more likely to divert their time to other activities such as sports.

Mother's education (medu) is found to be insignificant when analyzed by os, with a p-value greater than 0.1, but under quantile regression we find it to be significant at 0.75, and thus it would be implausible to estimate it using general os. In our study we found no significant direct effect of mother's education on scores, and indirect effects are not explored in depth in this paper.

The traveltime from home to school is found to be significant when analyzed by los with a p-value of less than 0.05, but it is not clear at which quartile it is significant, from the quantile results we know that it is significant at the 0.5 quartile. The time from home to school is clearly negatively significant, the longer the time to school the higher the grades, it could be that the parents expect the students to go to a better school and therefore it will show in the grades.

Weekly study time (STUDYTIME) is found to be significant if analyzed by ols with a p-value of less than 0.05, but it is not clear at which quartile it is significant, but by the quartile results we know that it is significant at the 0.25 quartile. The reason for this is that the students want to learn more so they will spend more time exploring the study.

The factor of wanting higher education (higher) is found to be significant when analyzed by os, but the significance varies at different quartiles, as can be seen in Table 2, the factor of enjoyment of higher education is not significant at the 0.25 quartile in affecting final grades, but it is indeed significant at the 0.5 and 0.75 quartiles. The coefficients of 2.43 at 0.5 and 2.72 at 0.75 are higher than the other variables, so the effect on the grades is very significant. The probable reason for this is that the students want to pursue higher education in the future and therefore will work harder in their studies.

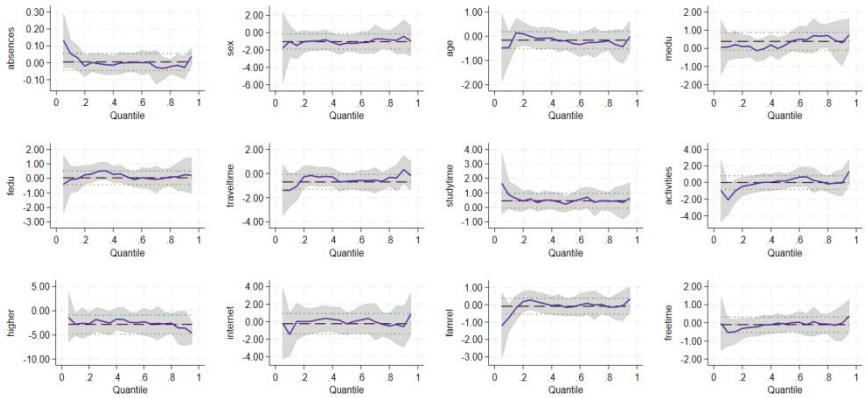

Figure 2 Regression analysis of student course quartiles in a secondary school

As can be seen in Figure 2，The amount of change in the number of absences is scored based on the number of absences, a low number of absences results in a high score, and a high number of absences results in a low score; a high number of absences directly affects a student's overall performance and future development. From the figure, it can be seen that in the case of low scores if the least square regression model is used to estimate the phenomenon of underestimation, the number of absences in the case of low scores significantly affects the overall performance of the students. Therefore students need to be aware of the number of absences so that it does not affect the overall performance of the students. The amount of variation in the distance from home to school data is based on the length of time scoring, the shorter the time, the lower the score, the longer the time, the higher the score; the longer the time the higher the score is because the school's quality of teaching and educational resources are better, the longer the distance students travel to school scores, the more the score is added to be considered a kind of incentive. The variable of extracurricular activities is based on the length of time of extracurricular activities, the longer the time of extracurricular activities, the more points, and the shorter the time, the less points. Extracurricular activities can contribute to the overall development of students. The variable of weekly study time is based on the data of weekly study time over a period of time (e.g., one month), as well as various factors that may affect study time. The study environment was found to have a significant effect on study time, and could be improved to enhance study efficiency. The variable of free time after school needs to be analyzed based on our need to clarify the nature of this variable and the question we want to explore. Free time after school is usually a continuous variable that indicates the period of time after school when students do not have scheduled activities or tasks, such as the amount of homework, participation in extracurricular activities, and the home environment. The greater the amount of homework, the less free time students have after school or the less free time students involved in extracurricular activities have after school.

4 CONCLUSION OF THE STUDY AND RECOMMENDATIONS FOR COUNTERMEASURES

In this paper, we use a combination of OLS and quantile regression to comparatively analyze the extent to which the factors affecting final grades of students receiving secondary education affect final grades and whether they are significant or not at different quartiles. The results of the analysis lead to the following conclusions: First, the variable of wanting to receive higher education significantly affects final grades, so students can be encouraged to aim for higher education in the future, which can improve their academic performance. Secondly, commuting time is also an important factor that affects the final grade, which depends on the parents' expectations of the students, so it is important for the teacher to communicate with the parents. Thirdly, the number of absences is also an obvious factor that affects students' academic performance, so teachers should strictly manage attendance in the classroom in order to improve students' academic performance. Fourthly, the gender factor shows that more female students have higher scores, which is due to the fact that male students have more activities that cause them to miss their homework. Therefore, the classroom teacher should plan the number of extracurricular activities or the time of the activities reasonably.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This study was supported by the Guangzhou Huashang College "Elderly Care Service Supply chain Service Quality and improvement path Research of Guangzhou" (2021HSQX11).

REFERENCES

[1] Luo Qiang, Yu Feifei. Gender Differences in Mathematics Achievement of Junior High School Students and the Mechanism of Influence of Individual Factors Based on the Data of Academic Quality Monitoring of Junior High School Students in the City of S. Journal of Mathematics Education, 2024, 33(2): 27-33.
[2] Yue, Q. Analyzing the Knowledge Mapping of the Current Research Field of Education Informatization in China. Journal of Engineering Studies. 2019, 11(4): 409-417.
[3] Fan Shuyuan Xiong Yonghong. Research on the Function of Teaching Forum in Open Laboratory Teaching. Experimental Technology and Management, 2010, 27(1): 26-28.
[4] Li Wan. Optimization of modeling teaching in the reform of mathematics teaching in colleges and universities under the concept of "1+X" . Learning Weekly, 2024, (31): 21-24.
[5] Deng WX, Zhang X, Chen JW, et al. Hot Spots and Trends of China's Vocational Education Informatization Research. Journal of Ningbo Institute of Vocational Technology, 2024, 28(1): 18-24.
[6] Liu Y. Innovative Application and Practical Exploration of Artificial Intelligence Technology in Higher Education Teaching Evaluation. Information Systems Engineering, 2024, (7): 165-168.
[7] Bie Dunrong. Theoretical Explanation of AI Technology Applied to Teaching and Learning in University Education. Teaching in Chinese Universities, 2024, (3/4): 39-44

[8] Kang Yueyuan, Zhang Junhong, Song Chunli. International Research on Mathematics Education in the Age of Mathematical Intelligence: Frontier Hot Spots and Future Prospects. Journal of Mathematics Education, 2024, 33(5): 67-73.
[9] Li Tieying, Wang Yunfeng. The Dilemma and Optimization Countermeasures of the Integrated Teaching Resource Base Construction of Civic and Political Science Classes in Universities, Secondary and Primary Schools. Zhenjiang Higher Education Journal, 2024, 37(4): 70-74.
[10] Song Qijie. Research on Digital Transformation and Teaching Reform of Secondary Vocational Education in Tourism. Foreign Trade and Economics, 2024, (10): 142-145.