International Journal of Economics and Finance; Vol. 6, No. 11; 2014
ISSN 1916-971X E-ISSN 1916-9728
Published by Canadian Center of Science and Education

Econometrics Analysis on Factors Affecting Student Achievement

Yueyi Sun[1]

[1] Department of Economics, Hong Kong Baptist University, Hong Kong, China

Correspondence: Yueyi Sun, Department of Economics, Hong Kong Baptist University, Hong Kong, China. E-mail: sunyueyi@126.com

Received: July 8, 2014 Accepted: August 22, 2014 Online Published: October 25, 2014

doi:10.5539/ijef.v6n11p1 URL: http://dx.doi.org/10.5539/ijef.v6n11p1

Abstract

Purpose: This study aims to identify school-level variables that influence academic outcomes, and to determine the extent of their influence. Using state-level panel data, this study estimates a simple achievement function to explore the nexus between three identified factors (*percentage of students eligible for reduced/free lunch program, school enrolment* and *per-pupil expenditure*) and student achievement (*percentage of satisfactory of 4th grade math and read*) in the United States.

Method: Based on literature reference and rational hypotheses, the effects of the *percentage of student eligible for reduced or free lunch, school enrolment* and *per-pupil expenditure* on the *percentage of 4th grade student satisfactory in math and read* were tested for a certain group of students separately. Ordinary Least Squares regression model was used to determine the validity and strength of each relationship.

Result: The data set consisted of 1823 observations located in different districts. Final test result shows that:

1) Significant negative effect on student achievement is found under the factor of percentage of students eligible for reduced or free lunch.
2) Slight negative effect on student achievement is found under the factor of school enrollment.
3) Slight positive effect on student achievement is found under the factor of per-pupil expenditure.

Students in school with lower percentage of students eligible for reduced or free lunch program, lower enrolment record, and more per-pupil expenditure will have better academic performance.

Keywords: student achievement, reduced or free lunch, school enrollment, per-pupil expenditure, social science

1. Introduction

What contributes to students' achievement is deeply concerned and hotly debated around the world. Researchers view this topic in various angles, both in traditional and contemporary ways. Some consider parental involvement as an important predictor on student achievement. Some argue that teacher characteristics matter more. While others claim that classroom and school factors relate strongly to student performance in mathematics. In this research, I set the focus on school quality and family financial background of students. This paper evaluates my assumptions by assessing the strength of each relationship between each independent variable and dependent variable.

For identifying variables, it's not difficult to define a good indicator for family financial background. Since the Free and Reduced Price Lunch program clearly presents income eligibility, it serves as a wonderful method to assess family financial background of students. The Free and Reduced Price Lunch program, included in the National School Lunch Program (NSLP), was established in 1946 under the National School Lunch Act. During the 2011–2012 school year, students in a family of four qualified for free lunch if their family income was less than $29,055. They qualified for the reduced rate if their families made less than $41,348.

However, a central issue of this debate lies in what factors constitute school quality. With scholars believing that small schools preserve individualized atmosphere and high teacher-student ratio, small class size and school size is considered a main contributor to high school quality. However, many critics argue that reducing class size only lead to a moderate gain in quality. Further explanation and interpretation of results will be elaborated in the RESULT part.

Moreover, there is a long-standing controversy whether improving school financial resources will promote

student performance. Per-pupil expenditure, as a general idea, needs to be specialized enough to determine its relationship with student achievement.

2. Literature Review

The main research issue concerns the factors that constitute school quality and family financial situation of students. The research hypotheses discussed in the following paragraphs are based on theoretical reasoning and results from previous studies. Given the importance of the issues examined in this study, we focus on the role of three variables to explain student achievement scores. Literature review is shown below:

The percentage of students participated in free/reduced-price lunch programs was considered a proxy for family financial situation as implied by Alan F. Meyers, Amy E. Sampson, Michael Weitzman, Beatrice L. Rogers and Herb Kayne (1989).

School enrolment also matters in predicting student achievement as identified by Holly Cato Bullard (2011). Some research indicates smaller schools facilitate higher achievement, and many other scholars verify this result. However, statistical analysis led researchers to conclude that no correlation existed between school enrollment and student performance in math or read. Because of the unclear relationship, two-tailed test was used later in testing, and I simply predict that the relationship is negative proved according to most theses.

Equally important in predicting student achievement is per-pupil expenditure. Hedges and Greenwald (1996) found either no or a weak relationship is between per-pupil expenditure and student achievement. Similarly, Kristen De Pena (2012) suggested that per-pupil expenditure has negligible effect on student performance, and Dennis J. Condron and Vincent J. Roscigno (2003) indicated that the partial effect of per-pupil expenditure on student achievement was very small. However, Childs and Shakeshaft (1986) concluded that per-pupil expenditure relating directly to instruction have the most positive influence on student achievement. Considering the lack of consistent findings, I take per-pupil expenditure as an independent factor, while assuming the partial effect on students' academic performance would be small.

Given these considerations, I formed the following hypotheses:

The family financial situation of students in a given school, measured by the percentage of students that participate in free/reduced-price lunch programs, will affect student achievement negatively.

School enrolment will affect student achievement negatively.

Per-pupil expenditures will affect academic achievement positively. However, the effect will possibly be very small.

3. Data Description

Carried with all literature referred above and all three hypotheses, relevant data was collected.

3.1 Dependent Variable y (Math4 or Read4)

State-wide assessment to measure achievement of students in public schools is having on record the achievement scores or percentage of students satisfactory of math and read. Cross-section data in terms of the percentage of 4^{th} grade students that reach the satisfactory level in mathematics achievement scores and reading achievement scores respectively was gathered, locating in different buildings and different districts.

Here I assume y equals to **math4** or **read4**, which stands for percentage of students satisfactory in 4^{th} grade mathematic or reading. The regression model runs twice using these two sets of data respectively, and it exposes structural similarity between the models of math4 and read4. However, the R^2 using **average math4 and read4** as dependent variable was lower than that of using math4 or read4 individually. Accordingly, I eliminated this approach of constructing the model.

3.2 Independent Variable x_1

Dataset	Mean	Standard Deviation
Lunch	39.25%	26.42

Correct form (proved by model specification test):

Lunch: percentage of students eligible for free or reduced lunch

Lunch can be a good proxy of parental income. According to the hypotheses, the percentage of students eligible for reduced or free lunch was investigated. Result showed that the mean of this group of data is 39.25% and the standard deviation is 26.42, implying a big variance among data. My research took this problem into account and

discussion will be elaborated in the following part.

3.3 Independent Variable x_2

Dataset	Mean	Standard Deviation
School enrolment	5.911	0.42
Correct form (proved by model specification test): *Lenroll:* logarithm form of school enrolment		

Higher school enrolment can result in less individualized atmosphere, lower teacher-student ratio and worse school climate. Thus, the school quality will be impaired if the class and school size is too big. Logarithm form of school enrolment (**lenroll**) is defined as an independent factor (x_2) to estimate student achievement.

3.3 Independent Variable x_3

Dataset	Mean	Standard Deviation
Total expenditure	8.533	0.215
Correct form (proved by model specification test): *Lexppp:* logarithm form of per-pupil expenditure		

Existing research substantiates the conclusion that expenditure on instruction and administration will have a positive effect on student performance because both result in reduced class size, which raises achievement score. However, the data doesn't specify the different dimensions of expenditure, which proves to be a restriction in interpreting the results.

To make it simple, expenditure per-pupil, referring to the total annual amount per student spent on all functions combined, was used in conducting the model, and it was calculated from total expenditure divided by school enrollment. According to model specification test, I define logarithm form of per-pupil expenditure (**lexppp**) as an independent factor (x_3) to estimate student achievement.

4. Methodology

Ordinary Least Squares regression model is used to determine the strength of each relationship. The proposed model is:

$$Math4= \beta_0+\beta_1 lunch+\beta_2 lenroll+\beta_3 lexppp+u$$

4.1 Test for Model Specification

A multiple regression model suffers from functional form misspecification when it does not account for the relationship between the dependent and independent variables properly. In this report, I have a systematic examination on the logarithms and quadratics form of explanatory variables.

4.1.1 Logarithmic Functional Form

Two models were tested to verify whether I should use Logarithmic functional form. First I try to use school enrolment (**enroll**) and per-pupil expenditure (**exppp**) as the independent variables x_2 and x_3. Second I replace by the logarithmic form of school enrolment (**lenroll)** and per-pupil expenditure (**lexppp)**.

$$Level\text{-}Level: Math4= \beta_0+\beta_1 lunch+\beta_2 enroll+ \beta_3 exppp+u$$

$$Level\text{-}Log: Math4= \beta_0+\beta_1 lunch+\beta_2 lenroll+\beta_3 lexppp+u$$

Using n=1823 observations in the data set, it is found that β_2 and β_3 is relatively small in the Level-Level Model, and the Adjusted R^2 is less than that in the Level-Log Model holding the explanatory variables constant. On the basis of scale of parameters and R^2, the Level-Log Model is preferred.

4.1.2 Models with Quadratics and Interaction Term

At this stage, we used Ramsey's (1969) **regression specification error test** (RESET) test to identify whether there is any misspecification in the general functions.

$$F_{2,1817} =1.441< F_{critic,\ 5\%}=3.84.$$

Thus, the proposed model is valid.

4.2 Other Tests

Table 1. Independent variable and their hypothesized effects on student achievement

Independent variable	Hypothesized Effect
Percentage of students eligible for free or reduced lunch	Negative
School enrolment	Negative
Per-pupil expenditure	Positive

Note. There is no unified conclusion for the effect of school enrolment on student achievement, so I choose the major one to follow.

Table 1 shows our predicted partial effects of each independent variable on the outcome variable generated from literature review. Following testing results were evaluated and compared with our hypothesized effect.

4.2.1 Test for Partial Effect of Each Variable on y: T-test

We used the *t* statistic to test whether a particular independent variable does have partial effect on the dependent variable. Table 2 illustrates the test result:

Table 2. Regression results of selected variables on % satisfactory in 4th grade math

Independent variables	β	t
Constant	46.19535	2.578613
Lunch	-0.471381	-32.95298
Lenroll	-4.839606	-5.273984
Lexppp	8.534118	4.647045
F-statistic=371.2903	R^2=0.379789	Adjusted R^2 = 0.378766

Note. $t_{critic,\ 5\%} = 1.96$, $t_{critic,\ 1\%} = 2.58$, $F_{critical,5\%}=2.60$.

Obviously, the three variables, lunch, lenroll and lexppp are all significant at 1% significance level, which are consistent with our prediction.

4.2.2 Test for Good-of-Fitness: R^2

The R^2 of the estimated model is 0.380, which means that lunch, lenroll and lexppp together explain 38.0% of the variation in student achievement in the data set. In terms of goodness-of-fitness, this estimated model explains the dependent variable very well.

4.2.3 Test for Overall Significance: F-Test

The resulting F-statistic is much bigger than critical value. Thus, all independent variables are jointly significant at 5% significance level. The variables in the estimated model do explain some variation in student academic achievement.

4.2.4 Test for Multi-Colinearity

We checked the value of the correlation coefficient between independent variables . Table 3 illustrates the test result:

Table 3. Correlation between selected variables and % satisfactory in 4th grade math

Variable	Lunch	Lenroll	Lexppp
Lunch	1.000000	-0.067514	0.221953
Lenroll	-0.067514	1.000000	-0.296612
Lexppp	0.221953	-0.296612	1.000000

Obviously there is no perfect linear relationship in the model (Rule of Thumb r>0.85−0.9). No multi-colinearity exists in this model.

4.2.5 Test for Heteroskedasticity: White Test

White Test is used to test for heteroskedasticity in the proposed model.

The test result is $F_{9,1813}$=34.58829 and it is much bigger than the critical value $F_{critic,5\%}$=3.10. Thus, heteroskdasticity is shown in the model we proposed. Possible reasons are followed:

(1) The variance of the data distribution of lunch is huge. However, after applying data segmentation and running the White Test, I found that heteroskedasticity still existed. Thus I presume that some information inherent in the data set is not included in the model.

(2) The data size is limited. Therefore, we cannot fully demonstrate the relationship between variables.

5. Results

Based on the above test results, we finally get the observed model.

$$\widehat{math4} == 40.729 - 0.467\widehat{lunch} - 4.690\widehat{lenroll} + 8.357\widehat{lexppp}$$

5.1 Lunch

As predicted, the results of regression indicate that the percentage of students eligible for free or reduced lunch has a negative effect on percentage of 4th grade students satisfactory in math at 1% significance level. 1% increase in the amount of students eligible for free or reduced lunch is estimated to lead to 0.471% decrease in4th grade math satisfactory rate. Lunch, a proxy of family financial situation, demonstrates an inverse relation with school performance. We can reach the conclusion that students from low-income families scored lower than students from high-income families did.

According to Comfort O. Okpala, Amon O. Okpala and Frederick E. Smith (2001), the reasons may lie in the lack of educational resource materials at home and academically supportive home environment in low-income households.

However, the huge range of 100% and the standard deviation of around 26.42% in the data distribution of percentage of students eligible for free or reduced lunch caught my attention. Thus I made a bold hypothesis that school enrolment and expenditure structure could have different directions of effects among schools with students from different family financial background, which means segmentation is highly needed.

In order to detect the existence of such possibility, I divided the data into three groups-the percentage of students eligible for free or reduced lunch of less than 15% (one standard deviation lower than mean), between 15% and 65% and more than 65% (one standard deviation higher than mean), namely high-income, medium-income and low-income family groups. I then ran the t test in each group to test the partial effects of each independent variable on **math4**.

Table 4 illustrates the test result:

Table 4. Specific model for low-income, middle-income and high-income schools

	High-income schools		Middle-income schools		Low-income Schools	
	β	t	β	t	β	t
Constant	61.93424	***1.127619***	16.65324	***0.714995***	28.13143	***1.085110***
Lunch	-0.358424	-2.787403	-0.357219	-10.84078	-0.531565	-4.985426
Lenroll	-6.342181	-2.536293	-4.542490	-3.936631	-0.064824	***-0.041877***
Lexppp	6.540191	***1.125529***	11.34674	4.806523	7.384465	3.180712
R^2	0.103691		0.135051		0.036680	

Note. $t_{critic,\ 5\%}$ = 1.96, $t_{critic,\ 1\%}$ = 2.58, the results in italic type are insignificant.

According to the test results, we can conclude that the partial effect of **lexppp** on **math4** is not significant at 1% significance level in the high-income group. And the partial effect of **lenroll** on **math4** is also not significant at 1% significance level in the low-income group.

Besides, the R^2 is 10.37%, 13.5% and 3.67% respectively in each group, which are too low to construct an effective model. Therefore, the idea of grouping is not validate.

Whereas the idea of grouping is rejected, the regression test in each group shows that both lunch and lenroll have a negative relationship with school performance while lexppp has a positive one. Such results are in line with my

prediction.

5.2 School Enrolment

School enrolment has a slightly negative effect on mathematics scores according to Table 2. This testing result is in accord with our literature review. An 1% increase in number students eligible for free or reduced lunch will lead to 0.04690% decrease in 4th grade math satisfactory rate.

William J. Fowler, Jr. and Herbert J. Walberg (1991) identified that keeping schools relatively small might be more efficacious and may exhibit rare consensus as a goal of educators, the public, and those seeking equality of opportunity for students. Also verified by Cotton and Kathleen (1996) is that, both the number and the varieties of extracurricular activities in which students participate are significantly higher in small schools than in large ones.

The rationale behind the results is that small schools have more individualized atmosphere, which contributes to better interpersonal relations between and among students, teachers and administrators. Teacher-student ratios, which in many states are based upon full-time equivalent (FTE) teachers, will surely be higher in small schools. This kind of school climate has a positive effect on school quality, and improves student achievement.

What contradicts to my expectation is the RESET test. Since researches indicated that there is an efficient scale, as demonstrated in Table 5, which means there is a diminishing strength of effect on student achievement as school enrollment becomes bigger. With this concern in mind, I then replaced lenroll by $enroll^2$ to reflect the existence of an efficient scale. However, it is proved insignificant by testing.

Table 5. Optimal school size recommendations – climate versus efficiency

Grade Level	Ideal Enrolment for Positive Climate and Order	Ideal Enrolment for Economic Efficiency
Elementary	300–400	450–700
Middle	300–600	600–800
High	400–800	800–1,200

Source: Safe Schools Facilities Planner: Improving School Climate and Order Through Facilities Design. North Carolina Department of Public Instruction, 1998.

Therefore, I presume that there still exists some limitation within the data of school enrollment.

Something worth mentioning is that, there are many opponents to the well-believed message that smaller class benefits all pupils. Clearly not every small school is terrific, since being small is not enough. The effort of reducing class size itself does not guarantee success without additional attention to teacher quality, increased funding, availability of necessary facilities, and community/district belief in the power of the reform.

5.3 Per-Pupil Expenditure

Based on the regression results illustrated in Table 2, we can identify that per-pupil expenditure correlates positively with mathematics scores, as proved by Verstegan, D. and King, R. (1998) and Bruce D. Baker (2012). A 1% increase in the percentage of students eligible for free or reduced lunch is estimated to lead to 0.08357% decrease in 4th grade math satisfactory rate. The test result is consistent with our prediction. Reasons behind how school enrolment affects student performance are identified below:

According to Harold Wenglinsky (1997), expenditures on instruction and the administration of school districts' central offices are positively related to class size, with more spending leading to more reduced size. Class size is, in turn, positively related to school social environment, with schools having more cohesive social environments when they have smaller classes. Finally, cohesive school social environments are positively related to students' achievement above and beyond students' social backgrounds. In other words, leading researchers in the area acknowledge that any effect of per-pupil expenditures on academic achievement depends on how the money is spent, not on how much money is spent.

I urge caution in interpreting the result since the data collected failed to distinguish among different types of spending. It's entirely possible that some spending patterns that create dead-end paths are involved in per-pupil expenditure. For example, the money can just as easily be spent on maintaining the same number of teachers, but at higher salary levels, without an essential increase in the quality of education.

This limitation of data explains why the result I tested slightly violates the conclusion reached by Coleman (1996), Hedges and Greenwald (1996) and William E. Bibb and Larry McNeal (2012), who found out that either

no relationship or a relationship that is weak or inconsistent is between per-pupil expenditure and student achievement.

5.4 Test for Read4

Aside from testing for math4, I also did the regression analysis for the *read4*. Table 6 illustrates the test result:

Table 6. Regression results of selected variables on % satisfactory in 4th grade read

Independent variables	β	t
Constant	35.26228	2.079992
Lunch	-0.462757	-34.18530
Lenroll	-4.540975	-5.229273
Lexppp	8.180469	4.707172
F-statistic=398.8599	R^2=0.396799	Adjusted R^2= 0.395804

Note. $t_{critic,\ 5\%} = 1.96$, $t_{critic,\ 1\%} = 2.58$, $F_{critical,5\%}=2.60$.

The observed model of *read4* is

$$\widehat{read4} == 35.262\text{-}0.463\widehat{lunch}\text{-}4.541\widehat{lenroll}+8.180\widehat{lexppp}$$

The test results of read4 are consistent with that of math4.

6. Conclusion

The main purpose of this research is to identify the factors affecting student achievement. Reduced or free lunch, school enrolment and per-pupil expenditure, which represent family income level and school quality respectively, were tested to be statistically significant in explaining the difference in 4th grade mathematics achievement scores, and the test on percentage of student satisfactory on 4th grade read showed consistent results with the one done on math4.

Combined with both math4 and read4 test, the result of regression analysis showed that *% of students eligible for free or reduced lunch* and *school enrolment* have negative effects on student achievement. However, *per-pupil expenditure* affects student academic performance positively. Among these three factors, the effectiveness of *% of students eligible for free or reduced lunch* is the largest, which implies that, keeping other factors constant, a school of students in relative worse family financial situation will result in poorer student achievements. These findings hold up to the hypotheses I made.

Moreover, I found no need to divide different income groups into segments. Also, there was no sign for an efficient school scale. These two findings violate the literature I referred and need to be further explored.

References

Alan, F. M., Amy, E. S., Michael, W., Beatrice, L. R., & Herb, K. (1989). School Breakfast Program and School Performance. *Am J Dis Child*, 1234–1239. http://dx.doi.org/10.1001/archpedi.1989.02150220142035

Anna, J. E., & Brian, K. (2013). *The Impact of School Size on Student Achievement: Evidence from Four States*. EDRE Working Paper No. 2013–03.

Barbara, F., Mark, S., & Steven, S. S. (1995). The Relationship between Parental Income and Academic Performance of Medical Students. *Academic Medicine, 70*(12), 1142–1144. http://dx.doi.org/10.1097/00001888-199512000-00019

Bates, J. T. (1993). Portrait of a Successful Rural Alternative School. *Rural Educator, 14*(3), 20–24.

Bradley, M., Gongshu, Z., Carolyn, C., Gary, W., Kris, K., Judy, W., & Helmuts, F. (2000). School Size and Relationship to Achievement and Behavior. *Public Schools of North Carolina*.

Bruce, D. B. (2012). Revisiting the Age-Old Question: Does Money Matter in Education? *The Albert Shanker Institute.*

Comforto, O., Amono, O., & Fredericke, S. (2001). Parental Involvement, Instructional Expenditures, Family Socioeconomic Attributes, and Student Achievement. *The Journal of Educational Research, 95*(2), 110–115. http://dx.doi.org/10.1080/00220670109596579

Cotton. & Kathleen. (1996). School size, school climate, and student performance.

Dennis, J. C., & Vincent, J. R. (n.d.). Disparities Within: Unequal Spending and Achievement in an Urban School District. *Sociology of Education, 76*(1), 18–36. http://dx.doi.org/10.2307/3090259

Eric, F. D., & Maria, F. I. (2014). Effects of poverty and quality of the home environment on changes in the academic and behavioral adjustment of elementary school-age children. *Journal of Clinical Child Psychology, 23*(4). http://dx.doi.org/10.1207/s15374424jccp2304_6

Harold, W. (1997). How Money Matters: The Effect of School District Spending on Academic Achievement. *Sociology of Education, 70*(2), 221–237. http://dx.doi.org/10.2307/2673210

Holly, C. B. (2011). *The Effects of School Enrollment Size on Student Achievement.* A Thesis submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University in partial fulfillment of the requirements for the degree of Master of Public Policy in Public Policy.

Ilyana, K. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review, 25*, 63–75. http://dx.doi.org/10.1016/j.econedurev.2004.10.003

Jeremy, E. F., Anna, R. H., & Ruth, N. L. T. (2013). Reducing School Mobility: A Randomized Trial of a Relationship-Building Invention. *Am Educ Res J.*

Kenneth, R. S. (2014). School Size and Its Relationship to Student Outcomes and School Climate. *National Clearinghouse for Educational Facilities.*

Kristen, D. P. (2012). Throwing Money At Education isn't Working. *State Budget Solutions.*

Lamb, S., & Fullarton, S. (2002). Classroom and school factors affecting mathematics achievement: a comparative study of Australia and the United States using TIMSS. *Australian Journal of Education.* http://dx.doi.org/10.1177/000494410204600205

Ramin, A., & Nabi, K. A. (2010). L2 Teacher Characteristics as Predictors of Students, Academic Achievement. *TESL-EJ.*

Verstegan, D., & King, R. (1998). The Relationship Between School Spending and Student Achievement: A Review and Analysis of 35 Years of Production Function Research. *Journal of Education Finance, 24*(2), 243–62.

Victor, B., Daniel, S., Dong-il, K., Marilyn, W., & Eric, S. (1995). Schools as Communities, Poverty Levels of Student Populations, and Students' Attitudes, Motives, and Performance: A Multilevel Analysis. *Am Educ Res J, 32*(3), 627–658. http://dx.doi.org/10.3102/00028312032003627

William, E. B., & Larry, M. (2012). The Relationship between Per Pupil Expenditures and Student Achievement in Tennessee. *American International Journal of Contemporary Research, 2*(3).

William, H. J. (2005). Parental Involvement and Student Achievement: A Meta-Analysis. *Harvard Family Research Project.*