Is Standard Test Scoring a Good Measure of Educational Performance? A Case Study of Public Schools in Connecticut

Lei Chen

School of Business, Jianghan University, Wuhan, Hubei 430056, China

lei.chen@jhun.edu.cn

Keywords: Standard test scoring, Public school system, Chow test.

Abstract. To investigate the relationship between students' performance on standard tests and school investment and potential family influence, we collected data for 110 towns or regional school districts in Connecticut and applied standard linear regression model to find out the most significant factors that may affect the test scoring. A chow-test was applied to check if there is a structural difference between the regional school district and the normal school district in each town. Since we used the cross-section data, a test for heteroscedasticity was applied. The result showed that the school investment, in terms of labor and capital inputs, was not important to the students' performance of standard tests, but the household income and parents' educational level seemed positively related to students' performance, and the percentage of non-English home language and percentage of low income families had a negative effect on the scoring.

Introduction

On January 8, 2002, United States President Bush signed into law the *No Child Left Behind* Act of 2001 (NCLB). According to the NCLB act, each state, school district, and school will be expected to make adequate yearly progress toward meeting state standards, which are measured by each student's performance of the standard tests of Math, Reading, Writing, and Science. This progress will be measured for all students by sorting test results for students who are economically disadvantaged, from racial or ethnic minority groups, have disabilities, or have limited English proficiency. In the state of Connecticut, for example, the Department of Education conducts standard Connecticut Mastery Test (CMT) for the 4th, 6th, and 8th grades and Connecticut Academic Performance Test (CAPT) for the 10th grade. There are three subjects (Math, Reading, and Writing) in the CMT test and four subjects (Math, Science, Reading, and Writing) in the CAPT test. Many researchers have showed their interests on this subject and done some research on it. For instance, Miron evaluated the performance of charter schools in Connecticut [1]. His research included 10 of the 14 currently operating charter schools in Connecticut, and used coherent analysis to investigate the progress of the performance of the involved schools by comparing the students' average CMT or CAPT score in 2001-2002 to that in 2003-2004. He found out that charter school students were gaining more on the state assessment tests than students in surrounding traditional public schools. Also, Vaz did an analysis of students' performance of CAPT in Connecticut technical high schools from 1999-2003 [2]. He applied multiple regressions, multivariate analyses of variance, and Scheffe's post hoc analysis on the data and found out that male students out-performed females in mathematics and science, but females excelled in reading and writing. He also found out that urban students' performance differed from suburban students, and Blacks, Whites, and Hispanics scored differently in mathematics and science. These results look interesting to me, so we collect data for public schools in Connecticut and try to use standard linear regression model to figure out what factors determine the public school students' performance of the tests most.

Data and Literature Review

To find out the relations between test scores and school investment and family influence, we collect data for the 4th, 6th, and 8th grade students' CMT test scores in Math, Reading, and Writing, and 10th grade students' CAPT tests cores in Math, Science, Reading, and Writing in 110 towns or regional school districts in 2004. We also collect the labor and capital inputs data such as classroom teachers per student, administrators per student, supporting staffs per student, square feet of academic building per student, and total expenditure per student on plant operation and maintenance, where the first three are labor inputs, and the last two are capital inputs. These data can be found at the Connecticut Department of Education website. Since family influence is another potential reason for performance variation, we collect data for the nine socio-economic factors in each town or regional school district as well, which include per capita income, median housing value, median household income, parents' educational level (measures as percentage of parents that have bachelor's degree or above), percentage of families below poverty, percentage of single-parent families, percentage of minority, percentage of non-English home-language families, and percentage of low income families. Among them, the first four are regarded as advantage factors and the last five are regarded as disadvantage factors, in the sense that the first four should be positively related to the test scores and the last five should be negatively related to the test scores. The data are available from the US Census 2000 Summary File 3. The selection of socio-economic factors is referenced from Ray [3] and Heffley [4].

Model and Result

To get the average of CMT and CAPT score we firstly calculated the average score of the three subjects for CMT for each grade and the average score of the four subjects for CAPT for 10th grade, and then calculated the average of the average scores of 4th, 6th, and 8th CMT and 10th CAPT. Initially I planed to use the average of CMT and CAPT score as the dependent variable, and run linear regression on all the school inputs and socio-economic factors. But I realized that there might have existed multicollinearity among the explanatory variables, so I used the simple rule of thumb [5] to check the multicollinearity by looking at the sample correlation coefficients between pairs of explanatory variables. If the correlation coefficient between two explanatory variables is greater than 0.9, we would judge that there exists a strong linear association and a potentially harmful collinear relationship. Under this criterion, we found as expected that there is multicollinearity between per capita income, median housing value, and median household income (Table 2). Since there is no other way to fix the problem, we simply eliminated per capita income and median housing value from my model. We run the regression of the average test score on thirteen independent variables for totally 110 towns or regional school district. The result is as the following (model I):

$$\begin{aligned}
ATS = {} & \underset{(21.78)}{252.10} - \underset{(-0.2973)}{0.52112\,TEA} - \underset{(-0.9533)}{9.1034\,ADM} - \underset{(-0.3423)}{1.7246\,STA} + \underset{(0.5017)}{0.0030\,POM} \\
& - \underset{(-0.4042)}{0.01251\,SQF} + \underset{(3.855)}{0.23043\,MHI} + \underset{(1.539)}{0.17915\,BDA} + \underset{(0.9466)}{0.32744\,POV} - \underset{(-0.1916)}{0.0336\,SPF} \\
& - \underset{(-1.799)}{0.16499\,MIN} - \underset{(-1.184)}{0.16959\,NEH} - \underset{(-2.340)}{0.26812\,LOI} \qquad \text{t-ratio}
\end{aligned} \tag{1}$$

where $R^2 = 0.7187$, number of observations = 110, and
ATS = Average test score
TEA = Teachers per 100 students
ADM = Administrator per 100 students
STA = Staff per 100 students
POM = Plant operation and maintenance expenditure per student
SQF = Square feet of building per student
MHI = Median household income
BDA = % of bachelor's degree or above
SPF = % of single-parent families

MIN = % of minority
NEH = % of non-English home-language
LOI = % of low income families

From the above result we can see that some explanatory variables have “wrong” signs from our expectation, i.e. all the labor inputs, square-feet of building per student, and percentage of families below poverty. Also, we notice that some of the variables are not significant (absolute value of t-ratio is less than 1.3), surprisingly including all the school inputs (labor and capital). This indicates that students’ performance of the standard tests is not likely related to what schools would do. In other words, schools could not do much to improve students’ performance on the tests, but families could influence students more, especially the wealthiness, the ethnicity, and the parents’ educational level of the family. Therefore, we narrow down the model to five explanatory variables and run the ordinary linear regression again, and get the following result (model II):

$$\underset{(60.21)}{ATS = 244.18} + \underset{(4.152)}{0.22930\ MHI} + \underset{(1.427)}{0.13798\ BDA} - \underset{(-1.446)}{0.17057\ NEH} - \underset{(-4.732)}{0.38452\ LOI} \tag{2}$$

where $R^2 = 0.6983$ and values in above parentheses are t-ratios.

This new model tells us that average test score is positively related to the household income and parents’ educational level, which is consistent with our expectation because wealthy families may be more concerned about children’s studies and highly educated parents are more likely able to and willing to tutor their children at home, which helps the kids perform better in the tests. The result also demonstrates that higher the percentage of students that do not speak English at home, lower the average test score of these students. This is not surprising because two subjects of the CMT or CAPT test are Reading and Writing, which directly measure students’ masteries of English. And this result is consistent with Heffley’s conclusion.

Since the data include 102 towns and 8 regional school districts, it is natural to think of checking if there is significant difference between these two groups. Therefore we introduce a binary variable and revise the model (model III):

$$ATS_i = \beta_1 + \delta_1 {*} D_i + \beta_2 {*} MHI_i + \beta_3 {*} BDA_i + \beta_4 {*} NEH_i + \beta_5 {*} LOI_i + u_i \tag{3}$$

where D_i=1 if i=1,2,…,102, and D_i=0 if i=103,104,…,110.

Using SHAZAM to run the regression one can get the following result:

$$\underset{(52.42)}{ATS = 246.88} - \underset{(-1.123)}{3.4710\ D} + \underset{(4.143)}{0.22850\ MHI} + \underset{(1.533)}{0.14885\ BDA} - \underset{(-1.432)}{0.16874\ NEH} - \underset{(-4.552)}{0.37260\ LOI} \tag{4}$$

where R_2 = 0.7019 and number of observations = 110.

The result shows that whether students enroll in a school that belongs to a town or a regional school district does not have a significant effect on the students’ performance on the tests, because the absolute value of the t-ratio for the coefficient of the binary variable is less than the critical value. Next we conduct a Chow test to check the possible structural difference between the data for towns and regional school districts. The Chow test statistic is given by Greene [6]:

$$F = \frac{[SSE_3 - (SSE_1 + SSE_2)]/k}{(SSE_1 + SSE_2)/(n-2k)} \sim F_{k,n-2k}, \tag{5}$$

where SSE_1 is the sum of squared error for group 1 (with 102 observations of towns), SSE_2 is the sum of squared error for group 2 (with 8 observations of regional school districts), SSE_3 is the sum of squared error for all 110 observations with the dummy variable, and k=5, n=110. Using

SHAZAM to run three separate regressions we get SSE_1=7633, SSE_2=48.337, and SSE_3=7835. Thus substituting the values into the above formula we get $F = 0.40$. Since the 95% critical value for $F_{5,100}$ is 2.31, we fail to reject the null hypothesis and conclude that there is no structural difference between towns and regional school districts.

Furthermore, since we are dealing with cross-section data, we need to check the heteroscedasticity. Thus we conduct a White's test. First we use ordinary least squares of model I to get all the residuals e_i , where i=1, 2,…, 110. Since there are 15 variables in $x \otimes x$ including a constant term, the regression of the squared least squares residuals on these 15 variables produces $R^2 = 0.111$. The chi-squared statistic is therefore 110(0.111) = 12.21. Since the 95 percent critical value of chi-squared with (15–1) degrees of freedom is 23.68, the hypothesis of homoscedasticity is not rejected by the test. In other words, there exists no heteroscedasticity.

Conclusion

After applying the data of 110 towns or regional school districts in Connecticut to my model we found out that public schools can not do much to improve the students' performance of standard tests. Nonetheless, living environments and socio-economic factors in the neighborhood can affect students' behaviors, especially household's income, parents' educational level, and languages being spoken at home. Thus instead of relying mostly on teachers and administrators from school to help students improve their test scores to achieve the goal set by the NCLB act, we would better seek some help from parents. Unfortunately the Connecticut Department of Education can only manage some school-related issues, but not parents' socio-economic status, e.g., increasing parents' salaries or sending them to college. Therefore we would expect that students' average testing score would not improve much in the next several years because it takes time for those socio-economic factors to change (hopefully in a positive way). One possible extension of this work is to keep track the data for the next several years and then do a panel data analysis to find out the effects over time.

References

[1] Miron, Gary. Evaluating the performance of charter schools in Connecticut. A report for ConnCAN, the Evaluation Canter, Western Michigan Univrsity. (2005)

[2] Vaz, Alvin. An analysis of student performance in Connecticut technical high schools as measured by 2001 CAPT and 2003 NOCTI assessments. A PhD Dissertation, Department of Educational Leadership, Central Connecticut State University. (2006)

[3] Ray, Subhash. Resource-use efficiency in public schools: a study of Connecticut data. Management Science, Vol. 37, No. 12, page 1620-1628. (1991)

[4] Hellfey, Dennis. Do schools fail the mastery test? The Connecticut Economy, Vol. 5, No. 2, page 5. (1997)

[5] Griffiths, William, Carter Hill, and George Judge. Learning and Practicing Econometrics. 3rd Edition, New York: John Wiley and Sons. (1993)

[6] Greene, William. Econometric Analysis. 5th Edition. N.J.: Prentice Hall, Englewood Cliffs. (2003)