The SAT Anomaly: Does Average Class Size Impact Average SAT Score?

Robert Lyle II, Christopher Kitchens, Kylie Griffin

ECON 3161: Econometric Analysis

Dr. Shatakshee Dhongde

11-18-2022

Abstract: Because the SAT has the potential to set the trajectory of students' educational careers, the SAT is arguably the most important examination for most American high schoolers. Considering the importance of the SAT examination, we sought to examine whether there is a relationship between average SAT scores with average high school class size per public school district. Using Georgia county-level data, our findings indicate that there is a minute and statistically insignificant correlation between average class size and average SAT score. Our results do show, however, that other variables such as median household income, average expenditure per pupil, and average educational attainment do have statistically significant correlations with average SAT score.

I. Introduction

According to the standard economic analysis, additional educational achievement is understood to be an important causal factor in socioeconomic status. Thus, policymakers seeking to improve socioeconomic outcomes have focused on increasing the abilities and opportunities of high school students to attend prestigious undergraduate and graduate programs. This goal is certainly a reasonable one as it seems wise to equip the next generation with the highest possible quality of human capital. Since one of the most important influences on educational attainment is a student's high school performance, educational policy has been centered around increasing student performance through various policy measures, one of the most popular being average class size reduction.

While there are myriad measures of student performance, such as the National Assessment of Educational Progress (NAEP), a student's SAT score is particularly important because of the central role it plays in the college application and acceptance process. As such, our analysis uses average SAT score as a proxy for student performance.

While there obviously are factors other than average class size that affect SAT scores, namely household income, personality traits, general intelligence, and parents' level of educational attainment, these factors lie largely outside of the purview of policymakers. Since education officials focus on variables they can influence, such as average class size and average expenditure per pupil, our research aims to estimate the impact of these variables on SAT scores. Specifically, our analysis aims to project the impact of average high school class size, average expenditure per pupil, median household income, and average educational attainment on a county's average SAT score.

When considering the educational effects of average class size, one can use common reasoning to formulate a hypothesis about the effect of class size on student performance. For example, it would be reasonable to assume that smaller class sizes lead to improved student performance due to factors such as increased teacher-pupil interactions, fewer distractions from course material, and greater financial and non-financial resources per pupil. In accordance with common reasoning, we predict that there will be a strong negative correlation between average high school class size and average SAT score.

II. Literature Review

Due to the importance of education for individuals and society at large, much time and resources have been poured into studying the impact of various policy prescriptions on student performance. Much of the economic literature on the subject focuses attention on the undergraduate level, or when the high school level is addressed, they tend to use different measures of student performance. That being said, there has been research into the impact of various variables on SAT score.

The College Board, who administers the SAT every year, collects extensive demographic information on who sits for the exam as well as their test score. According to data included in the organization's 2013 Validity Sample, there is a noticeable correlation between household income, parent's level of educational attainment, and a student's SAT score (Beard and Marini 2018). From an economic perspective, this observation seems to make intuitive sense: Students with greater resources at their disposal should perform better than those with less, all things held equal. Further, we may presume that being raised in a poor household also means that parents have less leisure time to spend engaging in enriching activities with their children and taking an active role in their children's education. Economic research has indicated household income to be one of the greatest factors influencing a student's performance, precisely for these reasons (Klick 2000).

After Tennessee's Student/Teacher Achievement (STAR) Project in the 1980s, educational economists conducted extensive studies to determine whether average class size had an impact on student's educational attainment. Project STAR was an experiment conducted in Tennessee public schools which randomly assigned students across the state in grades kindergarten through 3rd in classes of different sizes. Specifically, 11,600 students from 79 schools across the state in grades kindergarten through third were randomly placed in either a small class (13-17 students), an average-sized class (22-25 students), or an average-sized class with a teacher's aid. At the conclusion of four years, all students were returned to average-sized classes (Kreuger and Whitmore 2000). Participating students graduated from the Tennessee public school systems in the spring of 1998 (Kreuger and Whitmore 2000).

It should be noted that the data sample displays some important traits that might impact researchers' analyses. Project STAR schools were located in areas of modestly higher childhood poverty rates and slightly lower educational attainment than the Tennessee average. Student performance, as measured by the ACT, was slightly lower for STAR students than for Tennessee students as a whole. Further, at the time of the study, the average expenditure per pupil in Tennessee was only three-quarters of the national average (Kreuger and Whitmore 2000).

As with the literature at large on average class size's effect on student performance, research on the study's results are mixed. One study conducted by economists Alan Kreuger and Diane Whitmore indicated that the educational interventions of the State of Tennessee increased the likelihood that students sat for either the ACT or SAT (Kreuger and Whitmore 2000). Further, Kruger and Whitmore's work noted a slight positive correlation between smaller class size and average ACT score (Kreuger and

Whitmore 2000). [SAT scores were converted into their ACT equivalents in this study because 40% of the sample sat for the ACT while only 6% sat for the SAT. Due to the high correlation between ACT and SAT, we deemed this study relevant to our hypothesis (Dorans 2014).] Interestingly, their work indicates the gains from smaller class size may be concentrated on smaller age groups, with a dissipating impact as students age.

Alternatively, other literature reaches a different conclusion. Particularly, some economists have concluded that longitudinal data over the latter half of the twentieth century indicate that positive correlation between class size reductions and average SAT score (Hanushek 2000). Additionally, Hanushek counters Kreuger's and Whitmore's conclusion of the positive impact of the Tennessee STAR Project. As such, Hanushek concludes that the STAR Project has "limited policy implications," and presents data that point toward the limited success of past classroom reduction efforts (Hanushek 2000).

Another central area of study in education economics is the effect of average expenditure per student on student performance. Naturally, average expenditure per pupil and classroom size will be correlated in some manner, albeit not perfectly as schools provide resources other than teachers to students. As with the literature on the effect of classroom reduction on student performance, the body of research on average expenditure per pupil is mixed.

Some literature using state SAT score averages and expenditure per student conclude that higher levels of expenditure are correlated with higher SAT scores (Ram 2004). Specifically, Ram's work indicates that both SAT math and verbal scores increase with increased expenditure per student, with the math score increasing by a greater proportion than the verbal component. Despite the positive correlation between spending and SAT score, he concludes that the correlation, while being statistically significant, is so slight as to require significant increases in expenditure to raise average test scores by any noticeable amount (Ram 2000).

In a similar study, an economist examined Pennsylvania Department of Education data on average expenditure per pupil and Pennsylvania System of School Assessment (PSSA) test, which measures the basic reading and math skills of the state's fifth, eighth, and eleventh graders (Klick 2000). Klick concludes that dollars spent per student are not particularly impactful on student performance. Additionally, what impact spending does appear to have on student performance seems to decline as a student grows older. On the other hand, Klick remarks, "The most clear-cut finding of this study is that dollars definitely do make a difference. But that difference is not necessarily made in the schools; instead it is made in the home. Poverty, without fail, proved to be a significant determinant of whether or not [sic] a student will succeed in school" (Klick 2000).

Our work is unique because it focuses on a relatively understudied cohort - high school students - using an economically important measure of data - SAT score - that has tangible impacts on a student's career.

III. Data

While school-level data would have been preferable, 2018-19 county-level is the most recent and readily available for all variables included in our models. Because Baker, Clay, Quitman, and Taliaferro Counties did not report a school district average SAT score, our analysis contains 156 observations, including 155 Georgia county school districts plus the Atlanta Public Schools System. We include the Atlanta Public Schools System in our analysis because this school district is one of the state's largest network of public schools. Additionally, the data from a number of city school districts were combined with their county's school districts as these schools display similar demographic characteristics to the other schools in their respective counties (Appendix K). To better approximate the *ceteris paribus* condition, we use average expenditure per pupil, median household income, and average level of educational attainment as control variables.

Figure 1: Summary Statistics

Variable	Source	Observations	Mean	Minimum	Maximum	Standard Deviation
Average Expenditure/ Pupil (*AvgExp*)	GOSA	160	$10,344.63	$7997.86	$28,312.35	$2247.92
Average Class Size (*Class*)	OESE	159	15.54	7.6	21.4	2.09
Average SAT Score (*SAT*)	GOSA	156	997.13	882	1148	56.64
Median Household Income (*MedInc*)	BEA	160	$38,818.57	$19,726	$95,148	$9774.75
Average Educational Attainment (*Educ*)	UGA Carl Vinson Institute of Government	160	83.46%	5.78%	67.72%	95.51%

First, our dependent variable is average SAT score. Our data for average SAT score has been obtained from the Georgia Governor's Office of Student Achievement (GOSA) for the 2018-19 school year. The distribution of district average SAT scores is approximately normal, with a minimum value of 882 found in Hancock County and a maximum value of 1148 attained in Walker County.

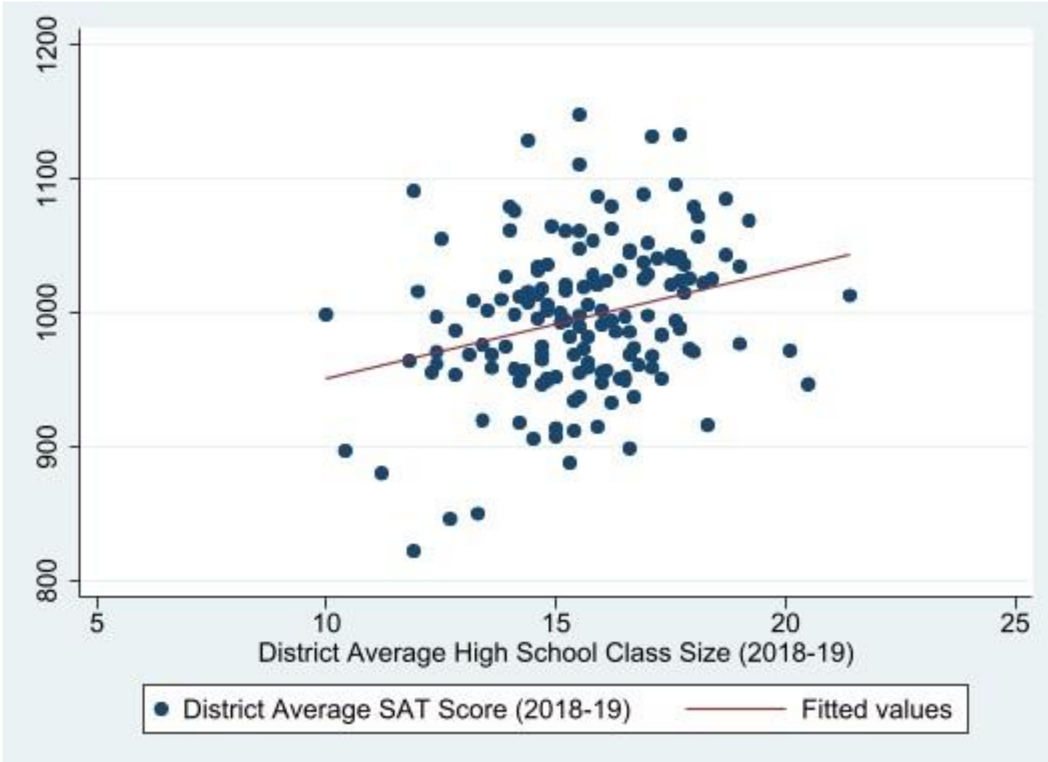
Our main independent variable is average class size. We use county-level average class size data from the Office of Elementary and Secondary Education (OESE) to tabulate average class size for the 2018-19 school year. Our analysis includes all high schools under the purview of the Georgia Department of Education, including traditional public schools, public charter schools, and magnet schools. Private and alternative high schools are not included in the data. District average class sizes range from 7.6 students per class in Taliaferro County to 21.4 students per class in Polk County. As a whole, this variable's distribution is approximately normal with a slight negative skew.

Our first control variable is average median household income per county in 2019. Our data is drawn from the Bureau of Economic Analysis (BEA). Median household incomes range from $19,726 in Walker County to $95,148 in Fulton County. Painting with broad strokes, the more sparsely-populated counties appear to fall on the lower end of the distribution, although there are notable exceptions. Median household income possesses an approximately normal distribution, albeit with a positive skew due to some outliers. The decision to include median income within our analysis is based on the robust correlation between SAT scores and income observed in related literature (Sockin 2021).

Our second control variable is average expenditure per pupil, measured according to data from GOSA for the 2019 fiscal year. Long County spent the least per pupil with an average expenditure of $7997.86 per pupil while Taliaferro County spent an average of $28,315.35 per pupil. The distribution of average expenditure per pupil possesses a noticeably positive skew due to a handful of counties significantly outspending their counterparts. We included this variable because it is a prominent subject in literature on educational attainment. Additionally, we predicted beforehand that average class size and average expenditures per pupil will be correlated in some way (Klick 2000).

Our final control variable is educational attainment. This variable measures the percent of a county's adult population that has a high school diploma or some higher degree. We use data from the University of Georgia's Carl Vinson Institute of Government. Atkinson County has the lowest percentage of educational attainment with a value of 67.8% of the adult population finishing at least high school. Fayette County has the highest educational attainment, with over 95% of adults finishing at least high school. Educational attainment is approximately normal, with a slight negative skew. Because of the common economic intuition that human capital development (*i.e.* education) is positively correlated with household income.

Figure 2: SLR Scatterplot



To assess the efficacy and efficiency of our model, we must determine whether our model meets the classic linear model assumptions. Examining the scatter plots between each of our regressors and average SAT score, we observe that there is a linear relationship between our estimates. Thus, we can reasonably assume that the relationship between these variables is similar in the population. Therefore, the first of the Gauss-Markov Assumptions is satisfied.

Next, our model meets the second Gauss-Markov Assumption of random sampling. Each of our observations come from different counties in Georgia such that each county's values are largely independent of each other.

Third, our model meets the non-perfect collinearity assumption. At first glance, none of our regressors appears to be a perfect linear combination of the others. Examining the variance inflation factors (VIF) of each regressor, we confirm this intuition as each VIF is well below the generally accepted cutoff value of 5 (Appendix H).

Fourth, the expected value of error does not depend on the values of our independent variables. More specifically, the expected error conditional on the expected values of our independent variables is equal to 0. This assertion is confirmed by examining the scatter between the residuals and fitted values (Appendix I).

On a similar note, our model meets the assumption of homoscedasticity. Examining the scatter plot between our simple linear regression model's error terms and residuals, our error term's variance does not appear to depend on the values of our regressors. Thus, we can reasonably assert that our model possesses a sufficient degree of homoscedasticity (Appendix I). Likewise, our model also meets the assumption of normality. By using a kernel density plot over a histogram of the residuals, we can see that the residuals approximate a normal distribution (Appendix J).

Since our model meets the Gauss Markov Assumption as well as the assumption of normality, we can state that our coefficient estimates are best linear unbiased estimators of the population parameters. Further, we may also assume that our estimators are normally distributed around the true parameters, which increases the validity of our confidence intervals and hypothesis tests.

IV. Results

To test our hypothesis, we use five regression models. We begin with the simple linear regression model with average class size regressing average SAT score:

Model 1: $SAT = \beta_0 + \beta_1\ (Class) + u$

Figure 3: Simple Linear Regression Model Estimates

$\boldsymbol{y}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$
SAT	870.236	8.120***

Note: Statistically Significant at 10%*, 5%**, 1%***

Model 1 possesses an R-squared value of 0.0776, which implies that average class size alone does not explain much of the variance of average SAT scores. It is also worth noting that the coefficient on *Class* is positive in a statistically-significant manner as this seems to contradict our initial hypothesis. Since this model does a poor job of simulating a *ceteris paribus* condition, we must introduce more variables that may have a relationship with average SAT score. Next, we show our most expansive model:

Model 2: $SAT = \beta_0 + \beta_1\ (Class) + \beta_2(AvgExp) + \beta_3\ (MedInc) + \beta_4(Educ) + \mathrm{u}$

Figure 4: Model 2 Estimates

$\boldsymbol{y}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$
SAT	868.388	-1.827	-0.014***	0.001***	294.653***

Note: Statistically Significant at 10%*, 5%**, 1%***

With an R-squared value of 0.3446, Model 2 fits our data much better. Further, the sign of the *Class* coefficient turns negative, resembling the relationship we initially predicted. However, the β_1 coefficient estimate is not statistically significant, so we do not have enough evidence in favor of our hypothesis. It is also interesting to note that the coefficient on *AvgExp* is also negative, which appears counter to what common intuition may posit. We define our third model by removing the independent variable – other than average class size – with the least statistical significance:

Model 3: $SAT = \beta_0 + \beta_1\ (Class) + \beta_2(AvgExp) + \beta_3(Educ) + \mathrm{u}$

Figure 5: Model 3 Estimates

$\boldsymbol{y}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$
SAT	753.660	-0.992	-0.011***	444.948***

Note: Statistically Significant at 10%*, 5%**, 1%***

Removing median household income produces some notable changes. First, the R-squared value of Model 3 drops to 0.3063, which is reasonable given that we removed a statistically significant variable. It is also important to note that our intercept value declined by a noticeable level. Next, we remove *Educ* from our model:

Model 4: $SAT = \beta_0 + \beta_1\ (Class) + \beta_2(AvgExp) + \mathrm{u}$

Figure 6: Model 4 Estimates

$\boldsymbol{y}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$
SAT	1051.021	3.266	-0.010***

Note: Statistically Significant at 10%*, 5%**, 1%***

In defining our fourth model, it is important note first that the sign of the coefficient on *Class* became positive, implying that increases in average high school class sizes contribute to an improvement in average SAT scores. Second, our R-squared values dropped precipitously to a value of 0.1204 while the value of our intercept increases by a value of almost 300. While this is not a particularly good model, as our regression estimates show, this model allows us to test the joint significance of median household income and educational attainment later in our analysis.

Model 5: $\ln(SAT) = \beta_0 + \beta_1(Class) + \beta_2(AvgExp) + \beta_3(MedInc) + \beta_4(Educ) + u$

Figure 7: Model 5 Estimates

$\boldsymbol{y}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$
$\ln(SAT)$	6.782	-0.002	-1.51e-05***	1.45e-06***	0.297***

Note: Statistically Significant at 10%*, 5%**, 1%***

In our final model, we apply a log-level functional form. This model has the highest R-squared value found in our analysis, signifying that this model possesses the greatest ability to predict district average SAT scores. We may also note that the relationship between average class size and percentage change in average SAT score has changed yet again. We may also observe that the relationship between the other variables and average SAT remains unaltered in any significant manner. According to this model, we may observe that of all our variables, educational attainment produces the greatest tangible effect on average SAT score, with every 1% increase in educational attainment contributing to a 0.297% increase in district average SAT score.

To test the null hypothesis that average class size does not have an effect on average SAT score, we must examine the corresponding p-values in our models. The model that presents the most evidence to reject the null hypothesis is our simple linear regression model, in which the p-value for *Class* is 0.000; however, since this model is weak, we must look to other models to better approximate the *Class* coefficient. In Models 2-5, our p-values for *Class* range from a least significant p-value of 0.705 in Model 3 to a p-value of .251 in Model 4. Thus, none of our stronger models present enough evidence for us to conclude that class is individually significant.

We can, with a reasonable level of confidence, conclude that the variables *MedInc*, *Educ*, and *AvgExp* are statistically significant at a 1% significance level. This evidence signifies that these variables do indeed have an impact on average SAT score. This conclusion is reinforced by the fact that each of our control variable's confidence intervals did not include 0 at a 5% level of significance.

Figure 8: Model Estimates Summary

	Model 1	Model 2	Model 3	Model 4	Model 5
Class	8.120*** (2.253)	-1.827 (2.565)	-0.992 (2.614)	3.266 (2.837)	-0.002 (0.003)
AvgExp		-0.014*** (0.003)	-0.011*** (0.003)	-0.010 (0.004)	-1.51e-05*** (3.51e-06)
MedInc		0.001*** (4.92e-4)			1.45e-06*** (4.95e-07)
Educ		294.653*** (84.760)	444.948*** (69.728)		0.297*** (0.085)
Intercept	870.236 (35.486)	868.388 (88.281)	753.660 (81.400)	1051.021 (74.906)	6.782 (0.089)
R-squared	0.0778	0.3446	0.3063	0.1204	0.3495

Note: Statistically Significant at 10%*, 5%**, 1%***

V. Extensions

To test the joint significance of our independent variables, we conduct two F-tests. First, we test the joint significance of *MedInc* and *Educ*. Our decision to select these variables is motivated by the strong positive correlation between the two (Appendix G). For this F-test, we define our restricted and unrestricted models using our Models 2 and 4:

Unrestricted: $SAT = \beta_0 + \beta_1\ (Class) + \beta_2(AvgExp) + \beta_3\ (MedInc) + \beta_4(Educ) + u$

Restricted: $SAT = \beta_0 + \beta_1\ (Class) + \beta_2(AvgExp) + u$

Using the F-statistic formula, we get the following:

$$\text{F} = \frac{(437{,}351.349\text{-}325{,}901.058)/(2)}{(325901.058)/(151)} = 25.819$$

Thus, our F-statistic is 25.819, which far exceeds the critical value of F = 3.00 necessary for us to conclude that *MedInc* and *Educ* are jointly significant at a 5% significance level. Thus, we may conclude that *MedInc* and *Educ* have both a joint and individual effect on average SAT score.

Next, we will conduct a second F-test to determine whether *Class* and *AvgExp* are jointly significant. We selected these two variables because of the strong negative correlation between them (Appendix G). For this F-test, we will use the following restricted and unrestricted models:

Unrestricted: $SAT = \beta_0 + \beta_1\,(Class) + \beta_2(AvgExp) + \beta_3\,(MedInc) + \beta_4(Educ) + u$

Restricted: $SAT = \beta_0 + \beta_1(MedInc) + \beta_2\,(Educ) + u$

Plugging our restricted and unrestricted sum of squared residuals as well as our restricted and unrestricted degrees of freedom into the F-statistic formula, we produce the following:

$$\mathrm{F} = \frac{(374{,}448.266\text{-}325{,}901.058)/(2)}{(325{,}901.058)/(151)} = 11.247$$

Thus, our F-statistic for this F-test is 11.247, which exceeds the corresponding critical value of 3.00 necessary for us to reject the null hypothesis that neither of these variables have an effect on average SAT score. However, we may not conclude that our independent variable of interest, *Class*, has an effect on SAT as the F-test is used to show that at least *one* of a subset of independent variables has an impact on the dependent variable.

Next, we chose to examine the impact a district's population may have on its average SAT score. We use a dummy variable, namely *AMed*, to indicate whether a county's population exceeds the median population for counties in Georgia. This method of modeling provides a simple way to represent counties as either relatively sparsely- or relatively densely-populated. We define our new model as follows:

Model 6: $ln(SAT) = \beta_0 + \beta_1(Class) + \beta_2(AvgExp) + \beta_3(MedInc) + \beta_4(Educ) + \delta_0(AMed) + \mathrm{u}$

Figure 9: Model 6 Estimates

$\boldsymbol{y}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$	$\boldsymbol{\beta_4}$	$\boldsymbol{\delta_0}$
$\ln(SAT)$	6.834	-0.004	-1.51e-05***	1.29e-06**	0.297***	0.017*

Note: Statistically Significant at 10%*, 5%**, 1%***

Thus, population appears to be a statistically significant factor at a 10% significance level. Since this model is a log-level model, we may interpret the coefficient on *AMed* as signifying that counties meeting our definition of "densely-populated" have a 0.017% higher average SAT score than a sparsely populated counties, holding all else equal. So, while population is a statistically significant factor at a 10% significance level, it does not seem to cause that great of a change in average SAT score in effect. Further, it is worth noting that the R-squared value for this model is 0.3624, making this model a somewhat better fit than our previous ones.

VI. Conclusion

In summary, we may not reject the null hypothesis that average class size has no effect on average SAT score as evidenced by the corresponding p-values produced by our analysis. However, we may conclude that median household income, average expenditures per pupil, and average educational attainment do have an impact on a public-school district's average SAT score as these variables had p-values reflecting statistical significance.

As such, our results concur with the existing body of literature as we failed to show that variance in average high school class size has a significant impact on average SAT score. One possible explanation for this phenomenon could be related to the fact that interventions in average class size are primarily impactful for younger students, with the impact waning as pupils mature and become more self-sufficient. Therefore, one possible direction for future research could be to examine whether average class size has an impact for pupils in elementary and middle school. Further, future treatments of this subject could benefit from including other district- and school-level data beyond state boundaries so as to study a more diverse sample of data.

References:

Beard, J., Marini, J. (2018). Validity of the SAT for Predicting First-Year Grades: 2013 Sat Validity Sample. *College Board*. https://eric.ed.gov/?id=ED582459.

Bennett, R. (1999). Using New Technology to Improve Assessment. *Research Reports,* Report 62, (March 1999). https://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.1999.tb01804.x.

Bureau of Economic Analysis. (2022). Personal Income by County, Metro, and Other Areas, 2021 [Data set]. https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas

Governor's Office of Student Achievement for the State of Georgia. (2019). 2019 Financial Efficiency Star Rating [Data set]. https://gosa.georgia.gov/accountability/financial-efficiency-star-rating/financial-efficiency-star-rating-archives.

Governor's Office of Student Achievement for the State of Georgia. (2019). 2018-19 SAT Scores (Recent) [Data set]. https://gosa.georgia.gov/dashboards-data-report-card/downloadable-data.

Hanushek, E. (2000). "Evidence, Politics, and the Class Size Debate." https://www.semanticscholar.org/paper/Evidence,-Politics,-and-the-Class-Size-Debate-Hanushek/ed9cbd047856e34b6e26ee2784d2480482a8d013.

Klick, J. (2000). Do Dollars Make a Difference? The Relationship between Expenditures and Test Scores in Pennsylvania's Public Schools. *The American Economist,* Vol. 44, No.1 (Spring, 2000), pp. 81-87. https://www.jstor.org/stable/25604191?origin=JSTOR-pdf.

Krueger, A., Whitmore, D. (2000). The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project Star. *NBER*. https://www.nber.org/papers/w7656.

Office of Elementary and Secondary Education. (2019). 2018-19 Per Pupil Expenditures: Georgia [Data set]. https://oese.ed.gov/ppe/georgia/.

Paulson, M. (2021). Is Income Implicit in Measures of Student Ability?. *Penn Wharton Budget Model*. https://budgetmodel.wharton.upenn.edu/issues/2021/9/28/is-income-implicit-in-measures-of-student-ability#:~:text=SAT%20math%20and%20ACT%20scores,ranging%20from%200.06%20to%200.10.

Ram, R. (2010). School Expenditures and Student Achievement: Evidence for the United States. *Taylor & Francis*, https://www.tandfonline.com/doi/full/10.1080/0964529042000239177.

University of Georgia Carl Vinson Institute of Government (2022). 2016-20 Georgia Education Data [Data set]. https://georgiadata.org/data/data-tables.

University of Georgia Carl Vinson Institute of Government (2022). 2016-20 Georgia Population Data [Data set]. https://georgiadata.org/data/data-tables

Appendix

A. Model 1 Output

```
. regress SAT Class

      Source |       SS           df       MS      Number of obs   =       156
-------------+----------------------------------   F(1, 154)       =     12.98
       Model |  38660.9332         1  38660.9332   Prob > F        =    0.0004
    Residual |  458574.503       154  2977.75651   R-squared       =    0.0778
-------------+----------------------------------   Adj R-squared   =    0.0718
       Total |  497235.436       155  3207.97055   Root MSE        =    54.569

------------------------------------------------------------------------------
         SAT | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       Class |   8.119744   2.253465     3.60   0.000     3.668051    12.57144
       _cons |   870.2364   35.48614    24.52   0.000     800.1339    940.3388
------------------------------------------------------------------------------
```

B. Model 2 Output

```
. regress SAT Class AvgExp MedInc Educ

      Source |       SS           df       MS      Number of obs   =       156
-------------+----------------------------------   F(4, 151)       =     19.85
       Model |  171334.378         4  42833.5946   Prob > F        =    0.0000
    Residual |  325901.058       151  2158.28515   R-squared       =    0.3446
-------------+----------------------------------   Adj R-squared   =    0.3272
       Total |  497235.436       155  3207.97055   Root MSE        =    46.457

------------------------------------------------------------------------------
         SAT | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       Class |  -1.826866   2.564868    -0.71   0.477    -6.894529    3.240796
      AvgExp |  -.0144382   .0034932    -4.13   0.000      -.02134   -.0075364
      MedInc |   .0014604   .0004917     2.97   0.003      .000489    .0024318
        Educ |   294.6527   84.76028     3.48   0.001     127.1834    462.1219
       _cons |   868.3877   88.28117     9.84   0.000     693.9619    1042.814
```

C. Model 3 Output

```
. regress SAT Class AvgExp Educ
```

Source	SS	df	MS			
Model	152292.697	3	50764.2322	Number of obs	=	156
Residual	344942.739	152	2269.36013	F(3, 152)	=	22.37
				Prob > F	=	0.0000
Total	497235.436	155	3207.97055	R-squared	=	0.3063
				Adj R-squared	=	0.2926
				Root MSE	=	47.638

SAT	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Class	-.9923556	2.614213	-0.38	0.705	-6.157241	4.172529
AvgExp	-.0111843	.0034012	-3.29	0.001	-.0179042	-.0044645
Educ	444.9483	69.72765	6.38	0.000	307.1878	582.7088
_cons	753.6597	81.40004	9.26	0.000	592.8381	914.4813

D. Model 4 Output

```
. regress SAT Class AvgExp
```

Source	SS	df	MS			
Model	59884.0872	2	29942.0436	Number of obs	=	156
Residual	437351.349	153	2858.50555	F(2, 153)	=	10.47
				Prob > F	=	0.0001
Total	497235.436	155	3207.97055	R-squared	=	0.1204
				Adj R-squared	=	0.1089
				Root MSE	=	53.465

SAT	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Class	3.266442	2.836768	1.15	0.251	-2.337851	8.870734
AvgExp	-.0103945	.0038148	-2.72	0.007	-.0179309	-.0028581
_cons	1051.021	74.90579	14.03	0.000	903.0383	1199.005

E. Model 5 Output

```
. regress lnSAT Class AvgExp MedInc Educ

      Source |       SS           df       MS      Number of obs   =       156
-------------+----------------------------------   F(4, 151)       =     20.28
       Model |  .17718774         4  .044296935   Prob > F        =    0.0000
    Residual |  .329749813       151  .002183774   R-squared       =    0.3495
-------------+----------------------------------   Adj R-squared   =    0.3323
       Total |  .506937553       155  .003270565   Root MSE        =    .04673

------------------------------------------------------------------------------
       lnSAT | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       Class |  -.0019027     .00258    -0.74   0.462    -.0070002    .0031948
      AvgExp |  -.0000151   3.51e-06    -4.31   0.000    -.0000221   -8.20e-06
      MedInc |   1.45e-06   4.95e-07     2.94   0.004     4.76e-07    2.43e-06
        Educ |   .2965274   .0852593     3.48   0.001     .1280722    .4649827
       _cons |   6.781513   .0888009    76.37   0.000      6.60606    6.956965
------------------------------------------------------------------------------
```

F. Model 6 Output

```
. regress lnSAT Class AvgExp MedInc Educ AMed

      Source |       SS           df       MS      Number of obs   =       156
-------------+----------------------------------   F(5, 150)       =     17.05
       Model |  .183691435         5  .036738287   Prob > F        =    0.0000
    Residual |  .323246119       150  .002154974   R-squared       =    0.3624
-------------+----------------------------------   Adj R-squared   =    0.3411
       Total |  .506937553       155  .003270565   Root MSE        =    .04642

------------------------------------------------------------------------------
       lnSAT | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       Class |  -.0039747   .0028268    -1.41   0.162    -.0095602    .0016109
      AvgExp |  -.0000151   3.49e-06    -4.34   0.000     -.000022   -8.25e-06
      MedInc |   1.29e-06   5.01e-07     2.57   0.011     2.97e-07    2.28e-06
        Educ |   .2699969   .0860611     3.14   0.002     .0999483    .4400454
        AMed |   .0167871   .0096631     1.74   0.084    -.0023063    .0358806
       _cons |   6.833981   .0932404    73.29   0.000     6.649747    7.018215
------------------------------------------------------------------------------
```

G. Independent Variable Correlation Output

```
. corr SAT Class AvgExp MedInc Educ
(obs=156)
```

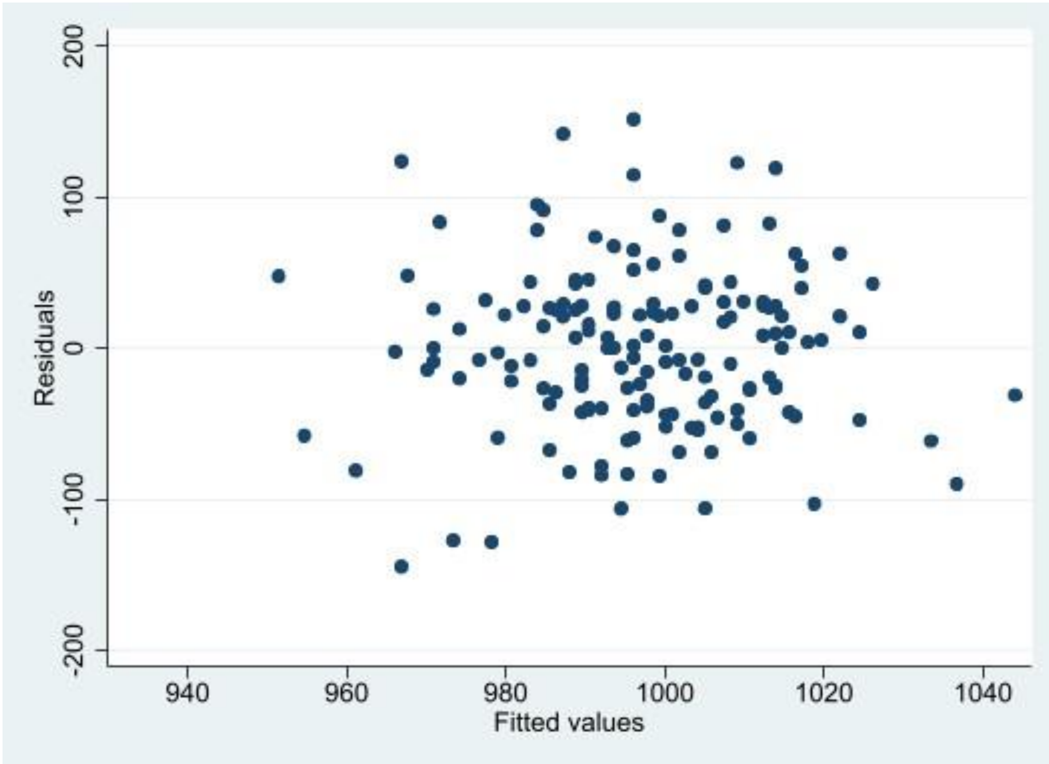
	SAT	Class	AvgExp	MedInc	Educ
SAT	1.0000				
Class	0.2788	1.0000			
AvgExp	-0.3359	-0.6279	1.0000		
MedInc	0.3637	0.0848	0.1606	1.0000	
Educ	0.4868	0.2958	-0.1587	0.5805	1.0000

H. Variance Inflation Factor Output

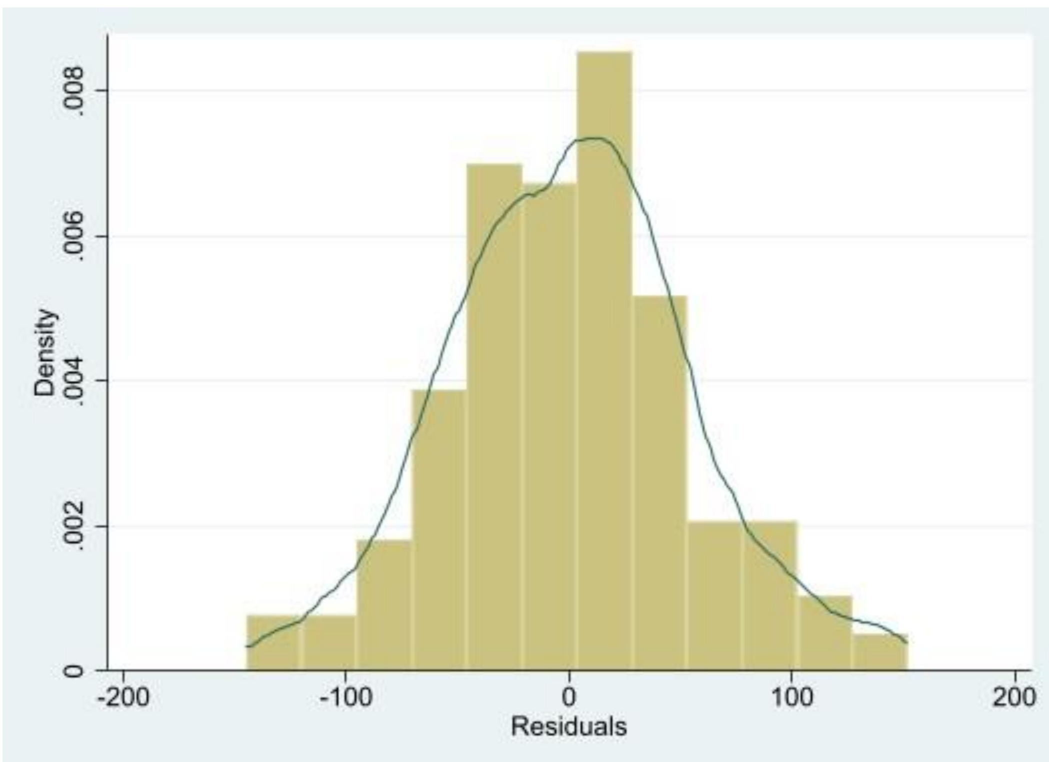
```
. vif
```

Variable	VIF	1/VIF
AvgExp	1.83	0.545466
Class	1.79	0.559489
Educ	1.70	0.586524
MedInc	1.69	0.590401
Mean VIF	1.75	

I. Homoskedasticity Graph



J. Distribution of Error Variance Graph



K. List of Georgia County Public School Districts

Appling, Atkinson, Atlanta Public Schools, Bacon, Baker, Baldwin, Banks, Barrow, Bartow, Ben Hill, Berrien, Bibb, Bleckley, Brantley, Brooks, Bryan, Bullouch, Burke, Butts, Calhoun, Camden, Candler, Carroll, Catoosa, Charlton, Savannah-Chatham, Chattahoochee, Chattooga, Cherokee, Clarke, Clay, Clayton, Clinch, Cobb, Coffee, Colquitt, Columbia, Cook, Coweta, Crawford, Crisp, Dade, Dawson, Decatur, DeKalb, Dodge, Dooly, Dougherty, Douglas, Early, Echols, Effingham, Elbert, Emanuel, Evans, Fannin, Fayette, Floyd, Forsyth, Franklin, Fulton, Gilmer, Glascock, Glynn, Gordon, Grady, Greene, Gwinnett, Habersham, Hall, Hancock, Haralson, Harris, Hart, Heard, Henry, Houston, Irwin, Jackson, Jasper, Jeff Davis, Jefferson, Jenkins, Johnson, Jones, Lamar, Lanier, Laurens, Lee, Liberty, Lincoln, Long, Lowndes, Lumpkin, Macon, Madison, Marion, McDuffie, McIntosh, Meriwhether, Miller, Mitchell, Monroe, Montgomery, Morgan, Murray, Muscogee, Newton, Oconee, Oglethorpe, Paulding, Peach, Pickens, Pierce, Pike, Polk, Pulaski, Putnam, Quitman, Rabun, Randolph, Richmond, Rockdale, Schley, Screven, Seminole, Griffin-Spalding, Stephens, Stewart, Sumter, Talbot, Taliaferro, Tattnall, Taylor, Telfair, Terrell, Thomas, Tift, Toombs, Towns, Treutlen, Troup, Turner, Twiggs, Union, Thomasville-Upton, Walker, Walton, Ware, Warren, Washington, Wayne, Webster, Wheeler, White, Whitfield, Wilcox, Wilkes, Wilkinson, Worth

Note that the data for following high school were included in the data for their respective counties:

Bremen High School – Haralson County; Buford High School – Gwinnett County; Calhoun High School – Gordon County; Cartersville High School – Bartow County; Carrolton High School – Carroll County; Chickamauga High School – Walker County; Commerce High School – Jackson County; Dalton High School – Whitfield County; Decatur High School – DeKalb County; Dublin High School – Laurens County; Gainesville High School – Hall County; Gordon Lee High School – Walker County; Jefferson High School – Jackson County; Marietta High School – Cobb County; Pelham High School – Mitchell County; Rome High School – Floyd County; Social Circle High School – Walton County; Thomasville High School – Thomas County; Trion High School – Chattooga; Valdosta High School – Lowndes County; Vidalia Comprehensive High school – Toombs County