Assessing the Effect of Reducing Class Size on Math Scores of the First Graders

1. Introduction

1.1 Background

Teachers, parents, and policymakers have been interested in reducing class size to improve the performance of students in their early-stage education in the United States due to its potential benefits to improve the attainment of the students in the short-term and long-term. The Student/Teacher Achievement Ratio (STAR) project attracted many researchers' attention because of its large sample size and the nature of the study design, which was funded by the Tennessee General Assembly and conducted in the late 1980s (Word, E.R., 1990). Many early and later follow-up researched showed positive results of attending small size classes. For instance, Pate-Bain (1992) reported that students attending the small classes made up 55% -78% of the top-scoring 10% from kindergarten through third grade, which indicates the positive and cumulative effect of small class size on the performance of K-3 graders. One follow-up research showed that attending smaller class size in primary school can increase the likelihood of earning a college degree by 1.6 percentage and the test-score effects were the excellent predictors for the long-term achievements (Dynarski, Hyman, & Schanzenbach, 2013).

To obtain more specific evidence of the effects of small class size on the performance of students, our project is to reassess the effects of reducing class size by examining the relationship between the class size and the math scores of students in the first grade using data from the STAR project.

1.2 Statistical Questions of Interest

The primary question of interest is whether the reduced class size associated with better students' math scores compared to large class sizes. To answer the primary question of interest, first, we need to know there are any extreme and missing values. Second, what are the characteristics and distributions of variables that may affect the results of the math test score for the first graders? Third, which model can fit our data? In addition, we need to verify our model through model diagnosis and/or sensitivity test.

2. Analysis Plan

2.1 Population and Study Design

The STAR study randomly assigned students and teachers into one of three level of class size that including small class (13 to 17 students per teacher), regular class (22 to 25 students per teacher), and regular-with-aide class (22 to 25 students per teacher), which started as the students entering school in kindergarten through the third grade; the schools enrolled in the study had at least one class of each type for proper randomization (Achilles et al, 2008).

We obtain the data from a package called "AER" in R. The data contains 11,598 observations on 47 variables including test scores of reading and math, the treatment groups (small class, regular class, regular class +aide), the characteristics of students (gender, ethnicity, birth quarter, and free-lunch status), the characteristics of teachers (the years of teaching experience, academic degree, career ladder, ethnicity), and the information of schools (Stock & Watson, 2007). For this project, we focus on the Math score of the first grader and the treatment groups.

2.2 Statistical Analysis

2.2.1 Exploratory Data Analysis

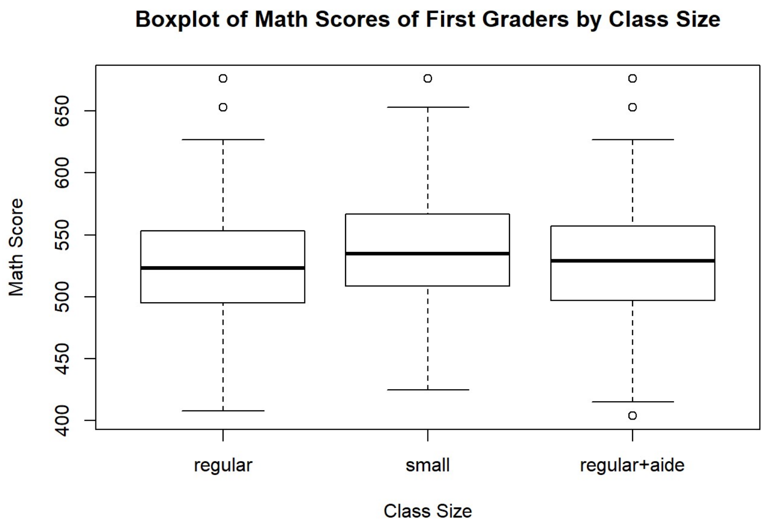
To find relative variables that reflect the association of difference of math scores of first graders, it is always a good idea to look at the data values and them summarize or graph them.

Firstly, we use the summary method in R to summarize the results of quantitative variables like math1(math scores in the first grade), read1 (reading scores in the first grade). In this way, we can see the maximum, minimum, median, quantile to check out whether we have extreme values or not.

Secondly, we explore the relationship between variables by boxplot. In this way, we can understand the characteristics of variables more intuitively and then choose suitable variables to analyze and fit our model.

Finally, since we want to test the effect of reducing class size on math scores of first graders, we mainly focus on two variables: math1(math scores of first graders) and star1(class size) and then remove the missing values using the list-wise deletion method by R package.

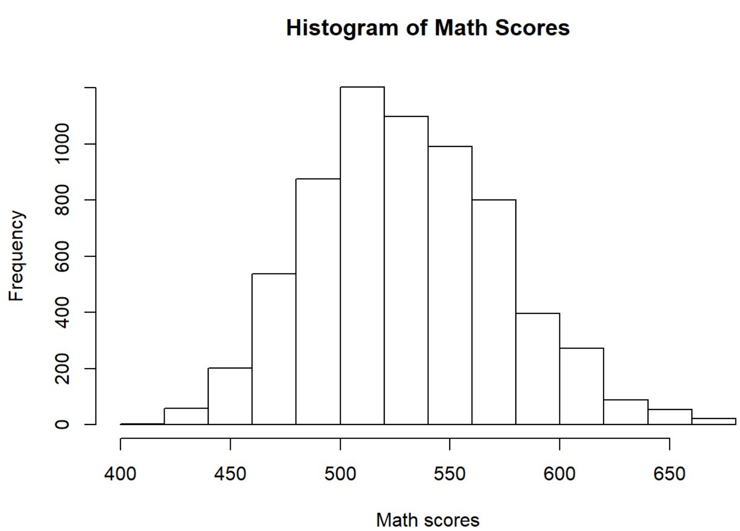
2.2.2 Statistical Modelling



From the boxplot of math scores of first graders by class size, it seems that the average math scores of 1st grade among three class size are different. The average score in the class of small size is the highest and the average score in the class of regular size is lowest. So, we choose these two variables: math1(math scores of first graders) and star1(class size) to analyze the effect of reducing class size on math scores of the first graders. In the case of randomized experiment design, one-way ANOVA model is suitable to interpret whether there is statistically significant difference of math scores in different class sizes.

In order to simplify the modeling methods, we make some basic Assumptions as follow:

1. The math scores in the first grade are independent with the scores in kindergarten, which means the math scores in kindergarten cannot influence the scores in the first grade.
2. The math scores of first graders cannot be influenced by gender, ethnicity, birth quarter, and free-lunch status. The randomized experiment design make sure that the math scores is associated with class size.



Through the preliminary analysis of data, the distribution of math scores of first graders seems to be bell-shaped. What’s more, the histograms of math scores in three class size have overlapped, which means that the variance of math scores in the different class size seems to be similar. Since the study randomly assigned students to small classes, regular classes, and regular classes with aide, it is reasonable to assume that the math scores in different class size are independent.

The assumptions of one-way ANOVA are shown as follow:

1. Normality: That each sample is taken from a normally distributed population;
2. Homoscedastic: That the variance of math scores in the different class groups should be the same;
3. Independence: that each sample has been drawn independently of the other samples.

Considering the following model notation:

Notation	Description

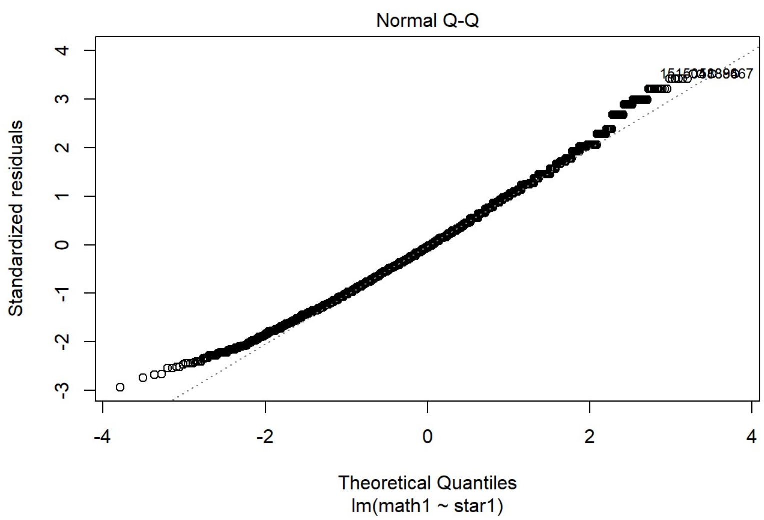
Notation	Description
i	Different class size: regular size if $i = 1$; small size if $i = 2$; regular size with aide if $i = 3$.
j	The index of math scores of the first graders, where $j = 1, 2, \ldots, n_i$.
n_i	The number of students in the i -th class, where $n_1 + n_2 + n_3 = n$ and n is the total number of students.
Y_{ij}	The j -th math scores of the first graders in the i -th class.
$x_{2,ij}$	The dummy variable that $x_{2,ij} = 1$ if the j-th math scores in the small size class($i = 2$); $x_{2,ij} = 0$ otherwise.
$x_{3,ij}$	The dummy variable that $x_{3,ij} = 1$ if the j-th math scores in the regular size with aide class($i = 3$); $x_{3,ij} = 0$ otherwise.
μ	The baseline intercept under the regular size class(as the reference group).
ϵ_{ij}	The random error which we cannot measure in the experiment.

The mathematical one-way ANOVA model with regression form is given by

$$Y_{ij} = \mu + \beta_2 x_{2,ij} + \beta_3 x_{3,ij} + \epsilon_{ij}$$

where $i = 1, 2, 3$, and β is parameters. In this experiment, the total number of students in first grades is $n = 6600$.

Then we use R to fit the one-way ANOVA model and test the difference of math scores in different class sizes.



Firstly, the original form of data are used to estimate the model parameters with least-square methods. After testing the model assumptions of original model, the normality cannot be satisfied with original model through the Q-Q plot since it seems that the distribution is relatively light-tailed in both sides. The Box-Cox transformation is needed to fix this problem, in order to maximize the log-likelihood in Box-Cox method, we apply log-transformation to the response variable Y_{ij} , the math scores. In the following step, we use $log(Y_{ij})$ as response variable to fit the one-way ANOVA model.

Secondly, when we fit the model, we can apply statistical testing methods to the data. To analyze the family-wise difference of math scores among each class size, the F-test is used with the one-way ANOVA model.

Finally, we use three pairwise comparisons among 3 class size groups to test pairwise effect of class size and to see which class size had the highest average math scores of first graders . In addition, one test based on the contrast of the difference between the math scores of small size class and the regular size class (regular size and regular size with aide) were conducted, in order and to compare the difference between small size and regular size. In this case, Bonferroni's procedure and Scheffe's procedure were both conducted in order to get more precise confidence intervals to do hypothesis testing. To test the data, 95% confidence intervals were derived to determine wheter the difference exists based on whether the interval contains zero.

2.2.3 Sensitivity Analysis

Since our model is based on some assumptions, in order to validate our model, we need to apply model diagnostics and sensitivity analysis.

In the assumptions of normality, we use Q-Q plot and Anderson-Darling normality test to check out whether the response variable is normally distributed. In the assumptions of homoscedastic, Residuals vs Fitted plot and Levene's test are used to test the variance of different class groups.

To obtain more consistent conclusion, we try to relax the assumptions of normality and homogeneity of variance of model. When relaxing the assumption of constant variance, Welch one-way test which is the for the pairwise comparisons with no assumption of equal variances can be used to further analysis. Moreover, Kruskal-Wallis rank sum test is a non-parametric method for testing whether samples originate from the same distribution, so we can apply non-parametric alternative to one-way ANOVA test with normality assumption.

3. Results

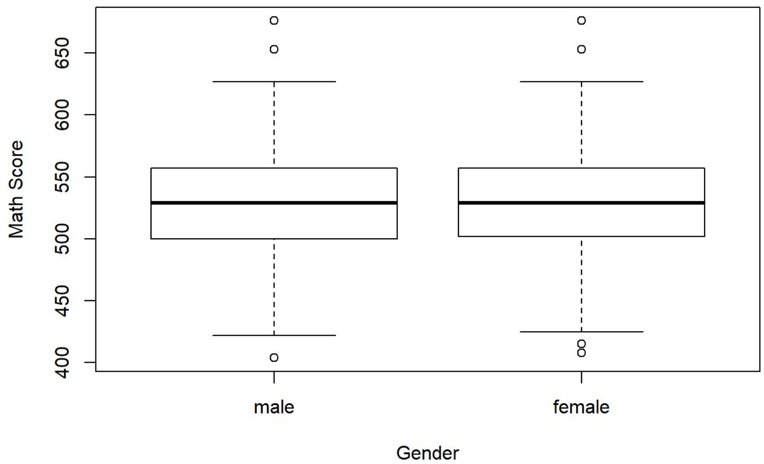
3.1 Exploratory Data Analysis

We check the summary statistics of every variables and find there is no abnormal value in the dataset. The proportion of missing values is relatively low, so we can use the data to analyze deeply. The detailed summary statistics are attached in the appendix.

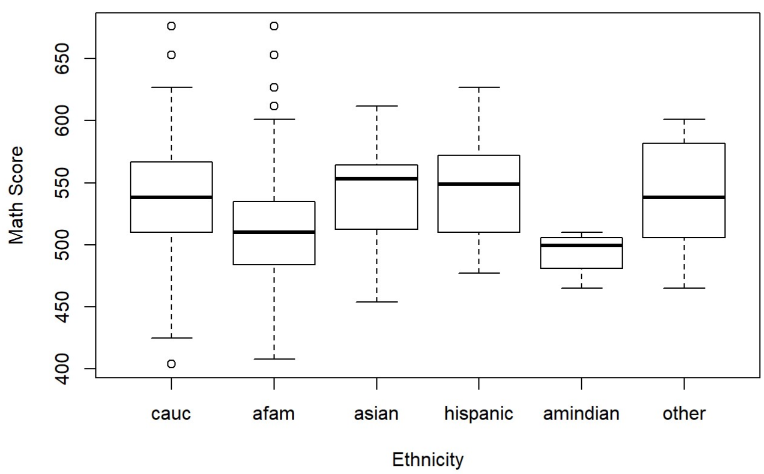
	regular (n=2584)	**small (n=1925)**	**regular+aide (n=2320)**	**Overall (n=11598)**
gender				
male	1326 (51.3%)	990 (51.4%)	1225 (52.8%)	6122 (52.8%)
female	1254 (48.5%)	934 (48.5%)	1087 (46.9%)	5456 (47.0%)
Missing	4 (0.2%)	1 (0.1%)	8 (0.3%)	20 (0.2%)
ethnicity				
cauc	1597 (61.8%)	1306 (67.8%)	1625 (70.0%)	7193 (62.0%)
afam	957 (37.0%)	599 (31.1%)	665 (28.7%)	4173 (36.0%)
asian	10 (0.4%)	7 (0.4%)	5 (0.2%)	32 (0.3%)
hispanic	2 (0.1%)	3 (0.2%)	4 (0.2%)	21 (0.2%)
amindian	3 (0.1%)	3 (0.2%)	3 (0.1%)	14 (0.1%)
other	4 (0.2%)	3 (0.2%)	4 (0.2%)	20 (0.2%)
Missing	11 (0.4%)	4 (0.2%)	14 (0.6%)	145 (1.3%)
read1				
Mean (SD)	513 (53.5)	530 (56.5)	521 (54.7)	521 (55.2)
Median [Min, Max]	507 [408, 651]	524 [417, 651]	516 [404, 651]	514 [404, 651]
Missing	148 (5.7%)	101 (5.2%)	184 (7.9%)	5202 (44.9%)
math1				
Mean (SD)	525 (41.7)	539 (44.1)	530 (42.9)	531 (43.1)
Median [Min, Max]	523 [408, 676]	535 [425, 676]	529 [404, 676]	529 [404, 676]
Missing	77 (3.0%)	57 (3.0%)	95 (4.1%)	4998 (43.1%)
lunch1				
non-free	1191 (46.1%)	967 (50.2%)	1063 (45.8%)	3221 (27.8%)
free	1341 (51.9%)	921 (47.8%)	1168 (50.3%)	3430 (29.6%)
Missing	52 (2.0%)	37 (1.9%)	89 (3.8%)	4947 (42.7%)
school1				
inner-city	570 (22.1%)	381 (19.8%)	429 (18.5%)	1380 (11.9%)
suburban	643 (24.9%)	452 (23.5%)	491 (21.2%)	1586 (13.7%)
rural	1163 (45.0%)	911 (47.3%)	1163 (50.1%)	3237 (27.9%)
urban	208 (8.0%)	181 (9.4%)	237 (10.2%)	626 (5.4%)
Missing	0 (0%)	0 (0%)	0 (0%)	4769 (41.1%)
degree1				
bachelor	1793 (69.4%)	1224 (63.6%)	1439 (62.0%)	4456 (38.4%)
master	791 (30.6%)	665 (34.5%)	838 (36.1%)	2294 (19.8%)
specialist	0 (0%)	17 (0.9%)	21 (0.9%)	38 (0.3%)
phd	0 (0%)	0 (0%)	22 (0.9%)	22 (0.2%)
Missing	0 (0%)	19 (1.0%)	0 (0%)	4788 (41.3%)
ladder1				
level1	1635 (63.3%)	1309 (68.0%)	1548 (66.7%)	4492 (38.7%)
level2	0 (0%)	15 (0.8%)	99 (4.3%)	114 (1.0%)
level3	89 (3.4%)	126 (6.5%)	76 (3.3%)	291 (2.5%)
apprentice	254 (9.8%)	203 (10.5%)	261 (11.2%)	718 (6.2%)
probation	382 (14.8%)	128 (6.6%)	156 (6.7%)	666 (5.7%)
notladder	201 (7.8%)	125 (6.5%)	180 (7.8%)	506 (4.4%)
Missing	23 (0.9%)	19 (1.0%)	0 (0%)	4811 (41.5%)
experience1				
Mean (SD)	10.3 (8.70)	12.2 (8.65)	12.7 (9.24)	11.6 (8.94)
Median [Min, Max]	8.00 [0.00, 42.0]	12.0 [0.00, 39.0]	11.0 [0.00, 39.0]	10.0 [0.00, 42.0]
Missing	0 (0%)	19 (1.0%)	0 (0%)	4788 (41.3%)
tethnicity1				
cauc	2150 (83.2%)	1556 (80.8%)	1886 (81.3%)	5592 (48.2%)
afam	418 (16.2%)	342 (17.8%)	424 (18.3%)	1184 (10.2%)

	regular (n=2584)	small (n=1925)	regular+aide (n=2320)	Overall (n=11598)
Missing	16 (0.6%)	27 (1.4%)	10 (0.4%)	4822 (41.6%)

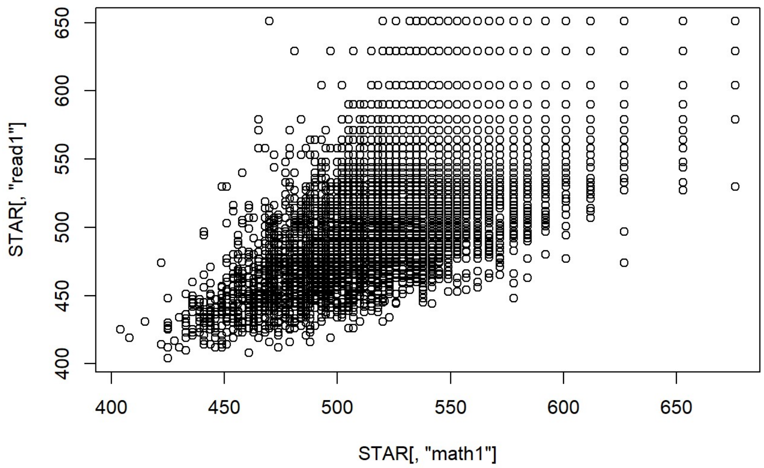
Since we are required to analyze the effect of different class size on math score in first grade. We assume that the scores in kindergarten have no influence on scores in first grade, so we mainly check out the summary statistics of the data in first grade. From the result,we can see that the quantitative variables don't have unreasonable data. Every type of variables contains low level of missing values, the data are suitable to analyze and have enough statistical power due to the big sample size. And then we can check out some relationships between each variables to choose suitable factors to fit our model.



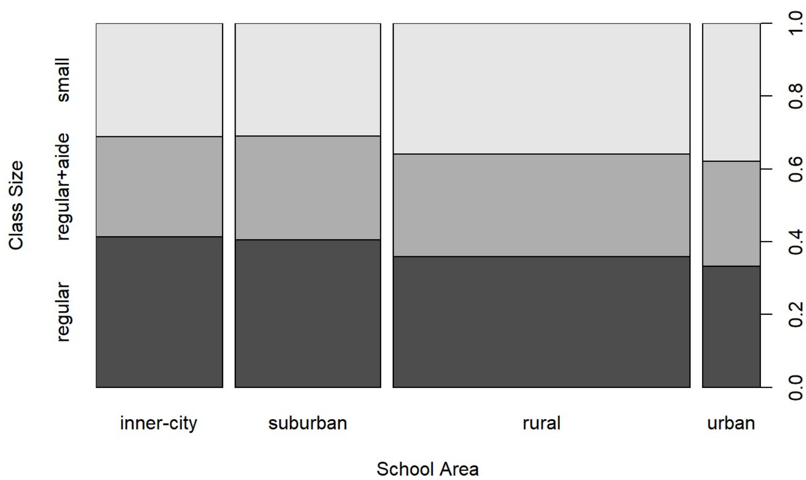
From the boxplot of math score by gender, we can see that gender seems has no influence on the math score since they hold a similar pattern between two groups.



From the boxplot of math score by ethnicity, we can see that the ethnicity does have influence on the math score, but maybe it is because the students have different educational background and there are other factors like economics and family background that influence the scores. Therefore we don't consider ethnicity as a predictor factor.

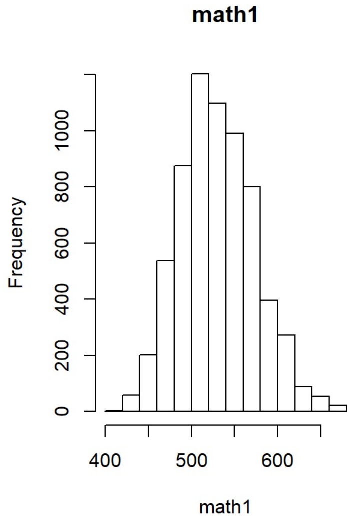


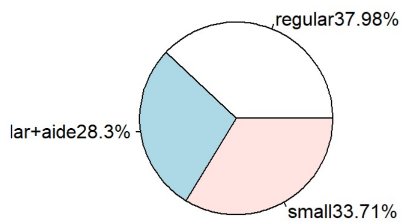
From the scatter plot, we can see that the reading score is positively correlated with the math scores. However, the math scores are what we focus on. We assume that the math score of the first grade can't be influenced by reading scores.



From this graph, we can find that the proportion of class size in different school areas are almost the same.

Base on the exploration above, we choose math1(math score of first graders) and star1 (class size) as our main research object. Since the data contains many missing values, we remove the missing values using list-wise deletion method by R package.



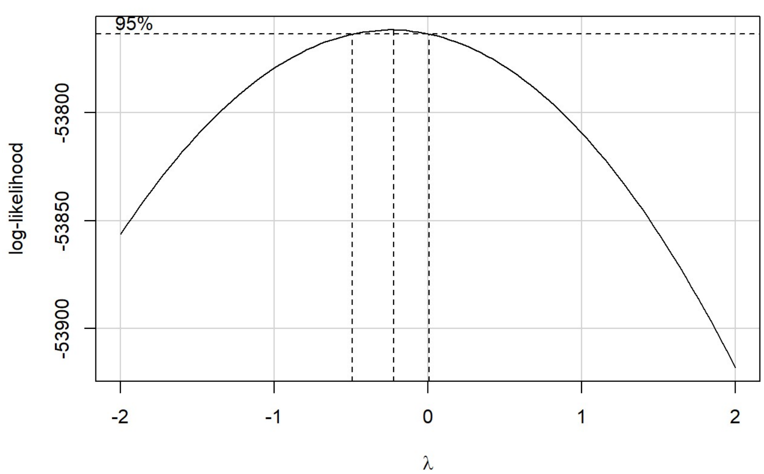


From the histogram, we can see that math1 is almost normal distribution and the class types are evenly distributed. The total sample size is 6600, which is big enough to hold statistical power and statistical inference.

3.2 Statistical Modelling

3.2.1 Transformation of Response

In order to fit a better model, we try to apply Box-Cox transformation to our response variable.



When we maximize the log-likelihood, the λ seems to be close to 0, which means that we should apply log-transformation to math score Y_{ij} . In the following step, we use $log(Y_{ij})$ as response variable to fit the one-way ANOVA model.

3.2.2 Model Fitting

```
##
## Call:
## lm(formula = log(math1) ~ star1, data = dat1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.267500 -0.056237 -0.001213  0.054563  0.255399
##
## Coefficients:
##                   Estimate Std. Error  t value Pr(>|t|)
## (Intercept)       6.260795   0.001604 3903.206  < 2e-16 ***
## star1small        0.024991   0.002455   10.181  < 2e-16 ***
## star1regular+aide 0.008121   0.002339    3.472 0.000521 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08031 on 6597 degrees of freedom
## Multiple R-squared:  0.01569,    Adjusted R-squared:  0.01539
## F-statistic: 52.57 on 2 and 6597 DF,  p-value: < 2.2e-16
```

The model summary provides most of the information we need to perform statistical inference. The fitted model form is:

$$log(Y_{ij}) = 6.260795 + 0.024991x_{2,ij} + 0.008121x_{3,ij}$$

From the t-test to each parameter, all estimated parameters in the model is statistically significant. The estimated group means(log-mean) are: $\hat{\mu}_1 = 6.260795$ for regular class; $\hat{\mu}_2 = 6.285796$ for small class; $\hat{\mu}_3 = 6.268916$ for regular class with aide.

3.2.3 Statistical Testing

Then, we can use ANOVA table and apply F-test to see family-wise difference of math scores in three class size.

Table 2 Results of ANOVA analysis

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Class size	2	0.6781075	0.3390537	52.5652	0
Residuals	6597	42.5516809	0.0064502		

AS shown in Table 1, class size had a significant effect on students' math score since p-value < 0.05, given the significant level. We are strongly believe that different class size have different impact on math scores of first graders.

Table 3 95% confidence intervals of multiple comparisons

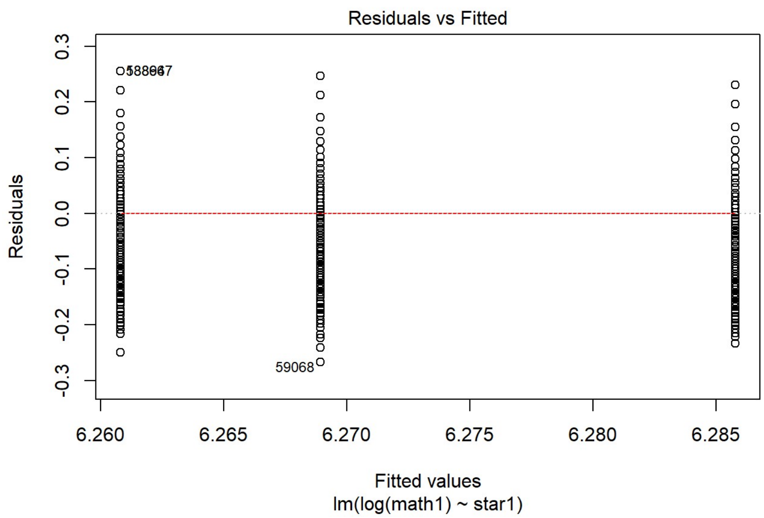
	Lower bound	Upper bound
Regular - Small	-0.0310008	-0.0189808
Regular - Regular with aide	-0.0138477	-0.0023936
Regular with aide - Small	-0.0230406	-0.0106997
Small - 0.5 (Regular + Regular with aide)	0.0155548	0.0263061

The results of multiple comparison (Table 3) showed that the students in small class tend to have the highest maths score, followed by the students in regular-with-aide class, and regular size class without aide had the lowest score. The mean math score of small size class is higher than the average math scores of regular size class with and without aide.

3.3 Sensitivity Analysis

The ANOVA test assumes that, the data are normally distributed and the variance across groups are homogeneous. We can check that with some diagnostic plots.

3.3.1 The Homogeneity of Variance

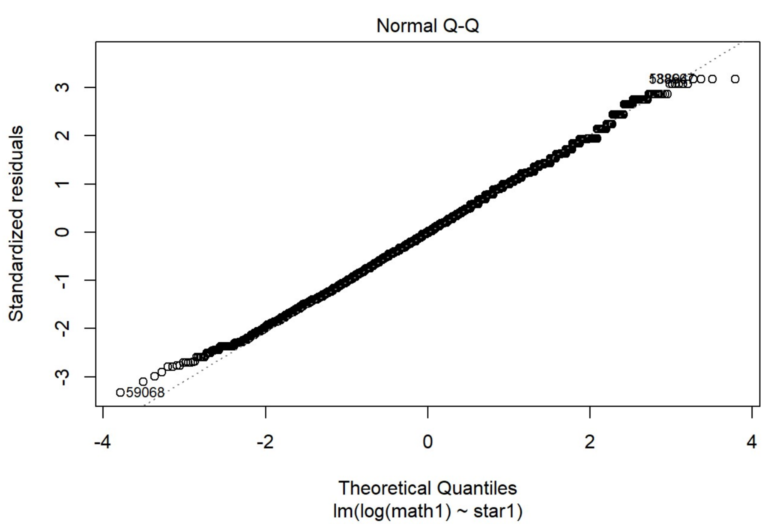


In the Residuals vs Fitted plot above, there is no evident relationships between residuals and fitted values (the mean of each groups), which is good. So, we can assume the homogeneity of variances.

```
## Levene's Test for Homogeneity of Variance (center = median)
##         Df F value Pr(>F)
## group    2  1.1836 0.3062
##       6597
```

From the output above we can see that the p-value is not less than the significance level of 0.05. This means that there is no evidence to suggest that the variance across groups is statistically significantly different. Therefore, we can assume the homogeneity of variances in the different kind of classes.

3.3.2 Normality Assumption



As all the points fall approximately along this reference line, we can assume normality.

```
## [1] 0.002017565
```

From the Anderson-Darling normality test, which works for larger sample sizes, the normality of random error terms is satisfied since p<0.05, given the significant level 0.05.

3.3.3 Relaxing the homogeneity of variance assumption

The classical one-way ANOVA test requires an assumption of equal variances for all groups. In our example, the homogeneity of variance assumption turned out to be fine since the Levene test is not significant. Welch one-way test, which is the ANOVA test with no assumption of equal variances can be used to further analysis.

Table 4 Results of Welch one-way test

F statistic	df of numerator	df of denominator	p-value
52.07045	2	4237.722	0

From the Welch one-way test, we can reject the null hypothesis that $\mu_1 = \mu_2 = \mu_3$ since $p < 0.05$, given the significant level 0.05.

Table 5 Results of pairwise comparisons adjusted by the Benjamini-Hochberg method

Small vs Regular	Regular vs Regular with aide	Small vs regular with aide
1.09532233861842e-23	0.000520713933634703	3.53276572558258e-11

We can also apply pairwise comparisons adjusted by the Benjamini-Hochberg method.The result is a table of p-values for the pairwise comparisons. Here, the p-values have been adjusted by the Benjamini-Hochberg method. Given the significant level 0.05, we can draw the same conclusions as the one-way ANOVA discussed above.

3.3.4 Non-parametric alternative to one-way ANOVA test

Table 6 Results of Kruskal-Wallis rank sum test

Chi-squared statistic	df	p-value
97.99319	2	0

Kruskal-Wallis rank sum test is a non-parametric method for testing whether samples originate from the same distribution. Given the significant level 0.05, we can draw the conclusions that math scores in 1st grade students are significantly different in the three classes with different size since $p < 0.05$.

4. Conclusion and Discussion

Our results showed that the small size class had a significant effect on 1st-year student's maths score, and students in small size classes tend to have higher maths score compared with regular class, which is consistent with other research by sensitive analysis. For instance, Folger, & Breda, (1989) reported that the score gains in small classes is around 15% higher than in regular classes for the first graders.The regular-size class with aide also made a difference to students compared with the students in regular-size classes. With the aide, students tend to have higher math scores.

However, we are unable to conclude that class size causes the increase of student's maths score because there are other factors, such as teacher's capability and school type, may play a more important role in student's test score. In addition, the model we built is based on some assumptions. We assume that the reading scores and math scores are independent but reading skills might affect our understanding of math formula in real life. Many mediators that might affect math scores cannot be eliminate if we relax our basic assumptions. Since the presence of missing values in the data, when we apply list-wise deletion, we might violate the randomized experiment design of the data. In conclusion, causality relationship cannot be directly achieved and more analysis like mediation analysis should be considered to deal with this situation. Further analysis of the data is required to examine the causal effect of class size on student's test scores.

5. Reference

Achilles, C. M., Bain, H. P., Bellott, F., Boyd-Zaharias, J., Finn, J., Folger, J., … & Word, E. (2008). Tennessee's Student Teacher Achievement Ratio (STAR) roject'. URL: http://hdl (http://hdl). handle. net/1902.1/10766.

Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. Journal of Policy Analysis and Management, 32(4), 692-717.

Folger, J., & Breda, C. (1989). Evidence from Project STAR about class size and student achievement. Peabody Journal of Education, 67(1), 17-33.

Pate-Bain, H., Achilles, C. M., Boyd-Zaharias, J., & McKenna, B. (1992). Class size does make a difference. Phi Delta Kappan, 74, 253-253

Stock, J. H., & Watson, M. W. (2007). Instrumental variable regression. Introduction to econometrics (2nd ed., pp. 421-467). Boston, MA: Pearson-Addison Wesley.

Word, E. R. (1990). The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Technical Report (1985-1990)

6. Appendix

```
## Warning in .table1.internal(x = x, labels = labels, groupspan = groupspan, :
## Table has 15 columns. Are you sure this is what you want?
```

	stark			
	1 (n=2194)	2 (n=1900)	3 (n=2231)	1 (n=258
factor(gender, labels = c("male", "female"))				
male	1119 (51.0%)	977 (51.4%)	1154 (51.7%)	1326 (51.
female	1075 (49.0%)	923 (48.6%)	1077 (48.3%)	1254 (48.
Missing	0 (0%)	0 (0%)	0 (0%)	4 (0.2%
factor(ethnicity, labels = c("cauc", "afam", "asian", "hispanic", "amindian", "other"))				

	stark			
	1 (n=2194)	2 (n=1900)	3 (n=2231)	1 (n=258
cauc	1472 (67.1%)	1294 (68.1%)	1468 (65.8%)	1597 (61.
afam	710 (32.4%)	593 (31.2%)	755 (33.8%)	957 (37.(
asian	8 (0.4%)	3 (0.2%)	3 (0.1%)	10 (0.4
hispanic	0 (0%)	4 (0.2%)	1 (0.0%)	2 (0.1%
amindian	1 (0.0%)	1 (0.1%)	0 (0%)	3 (0.1%
other	1 (0.0%)	4 (0.2%)	4 (0.2%)	4 (0.2%
Missing	2 (0.1%)	1 (0.1%)	0 (0%)	11 (0.4
as.numeric(reading_score)				
Mean (SD)	42.4 (17.3)	45.7 (17.5)	42.6 (16.8)	103 (29
Median [Min, Max]	41.0 [1.00, 93.0]	44.0 [5.00, 93.0]	40.0 [6.00, 93.0]	114 [26.0,
Missing	188 (8.6%)	161 (8.5%)	187 (8.4%)	148 (5.7
as.numeric(math_score)				
Mean (SD)	23.9 (7.59)	25.0 (7.52)	23.9 (7.21)	70.7 (18
Median [Min, Max]	24.0 [2.00, 38.0]	25.0 [1.00, 38.0]	24.0 [3.00, 38.0]	75.0 [16.0,
Missing	162 (7.4%)	138 (7.3%)	154 (6.9%)	77 (3.0
factor(lunch_type, labels = c("non-free", "free"))				
non-free	1143 (52.1%)	1001 (52.7%)	1105 (49.5%)	1191 (46.
free	1044 (47.6%)	891 (46.9%)	1117 (50.1%)	1341 (51.
Missing	7 (0.3%)	8 (0.4%)	9 (0.4%)	52 (2.0
factor(school_area, labels = c("inner-city", "suburban", "rural", "urban"))				
inner-city	507 (23.1%)	400 (21.1%)	521 (23.4%)	570 (22.
suburban	443 (20.2%)	456 (24.0%)	513 (23.0%)	643 (24.
rural	1074 (49.0%)	862 (45.4%)	981 (44.0%)	1163 (45.
urban	170 (7.7%)	182 (9.6%)	216 (9.7%)	208 (8.0
factor(degree_of_teacher, labels = c("bachelor", "master", "specialist", "phd"))				
bachelor	1406 (64.1%)	1304 (68.6%)	1409 (63.2%)	1793 (69.
master	715 (32.6%)	512 (26.9%)	754 (33.8%)	791 (30.(
specialist	0 (0%)	43 (2.3%)	0 (0%)	0 (0%
phd	73 (3.3%)	41 (2.2%)	47 (2.1%)	0 (0%
Missing	0 (0%)	0 (0%)	21 (0.9%)	0 (0%
factor(ladder, labels = c("level1", "level2", "level3", "apprentice", "probation", "notladder"))				
level1	1517 (69.1%)	1383 (72.8%)	1771 (79.4%)	1635 (63.
level2	46 (2.1%)	27 (1.4%)	46 (2.1%)	0 (0%
level3	41 (1.9%)	13 (0.7%)	0 (0%)	89 (3.4
apprentice	166 (7.6%)	190 (10.0%)	158 (7.1%)	254 (9.8
probation	176 (8.0%)	89 (4.7%)	69 (3.1%)	382 (14.
notladder	23 (1.0%)	14 (0.7%)	0 (0%)	201 (7.8
Missing	225 (10.3%)	184 (9.7%)	187 (8.4%)	23 (0.9
as.numeric(experience)				
Mean (SD)	12.0 (8.19)	12.9 (8.18)	12.5 (8.27)	13.7 (9.
Median [Min, Max]	9.00 [1.00, 25.0]	13.0 [1.00, 25.0]	10.0 [1.00, 25.0]	13.0 [1.00,
Missing	0 (0%)	0 (0%)	21 (0.9%)	0 (0%
factor(tethnicity, labels = c("cauc", "afam", "asian"))				
cauc	1737 (79.2%)	1630 (85.8%)	1865 (83.6%)	2150 (83.
afam	430 (19.6%)	264 (13.9%)	337 (15.1%)	418 (16.
asian	0 (0%)	0 (0%)	0 (0%)	0 (0%
Missing	27 (1.2%)	6 (0.3%)	29 (1.3%)	16 (0.6