What effect does the “quality” of a school have on the future economic prospects of those who attend it?

Fogofoluwa Afolabi

Registration Number: 1703675

Supervisor: Dr Elif Kubilay

EC831 Project: Economics

Date: 30th April 2021

Word Count: 7687

Table of Contents

1. Abstract

This paper examines the effect that school quality has on economic outcomes by using data from the Project Star experiment and supporting this with the empirical evidence in the literature. The purpose of the Project Star experiment was to use class size as a proxy for school quality to analyse the effects on students' academic test results. Project Star is a large scale experiment that aimed to show legislators the importance of small class sizes. The effect of school quality, in relation to class size is found to be positive but not necessarily significant when using the Ordinary Least Squares estimator. The findings suggest smaller classes raise the test scores of the students for the most part, but this is not always the case. The effects of other school quality variables like teacher student ratio and teacher experience are not conclusive.

2.Introduction

Schooling is a key pillar of most economies and is designed to prepare us all with the skills and subject specific knowledge needed in our adult lives. The project tackles the effects of school quality on future outcomes. This is particularly relevant as schooling is heterogeneous. Schools in the United Kingdom are categorised as either private or state. Within the state category, there is a branch called grammar schools, funded by the state that are deemed to be more selective than ordinary state schools. Generally, private schools offer a well-rounded school experience that goes far beyond the teaching. Public schools do not have the level of resources private schools have. From this, it can be said that the education system is a discriminatory process in that all children do not receive the same education. The research question at hand allows a review of this discrimination backed with empirical evidence.

One of the reasons why this question is necessary is that there are both short term and long-term consequences of school quality. Some of these short-term consequences are achievement in test scores and the dropout rate of pupils. Part of the variation in earnings can be explained by schooling quality and this is related to the long term. The quality of a school offers some insight into whether a student furthers their education or not. Partaking in schooling is both an investment and a cost simultaneously. The perception by individuals of how schools can contribute to their economic outcomes influences their saving patterns. Regardless of how people feel, our schooling achievements are our representative, hence are included in CVs. If the effects of school quality are purposeful in explaining future economic outcomes, this will encourage schools to prioritise quality of schooling not just quantity.

When research and policy makers come together it is often impactful. Policy makers look to research when implementing policies, especially since these policies affect society as a whole. If the research is meaningful, policy makers can create or develop policies centred around improving school quality. Afterwards, the government is able to allocate a share of their budget and provide support through resources and initiatives appropriately.

This question also brings to the forefront the inequality of income issue. Low-income families are more likely to live in areas where the schools are predominantly comprehensive and do not have the means financially to send their children to schools of higher quality. Good quality schooling should be accessible to all, irrespective of their background. Research like this may help in establishing the school environment in which minority students can excel best in.

Lastly, if the research accomplishes nothing else, it will put citizens in a better position to make informed decisions. Such a decision can be how much to invest in human capital. This research is also valuable for parents when making decisions on the schools their children should apply to.

The aim of this paper is to study what effect the quality of a school has on economic outcomes. The short term and long-term effects of this relationship will be addressed. In order to do this, the dataset from the Project Star experiment is employed. The empirical model of choice is the multiple linear regression model and the estimator is the Ordinary Least Squares. Multiple regressions are carried out where scores achieved in tests by the subjects is the dependent variable. A ceteris paribus approach is taken concerning the effects of class size which is achieved through control variables. The Breusch Pagan test was conducted and the outcome showed that heteroskedasticity

was present in the model. The appropriate action to use robust standard errors was taken. Overall, a critical analysis of the literature will be conducted.

Different pieces of literature use different indicators as a measurement for school quality. There is a lack of consensus among researchers regarding the effects of school quality on future prospects. Some economists argue that there are significant positive effects while others argue that the effects are insignificant.

Card and Krueger (1992) find school quality produces positive effects on students' future prospects, specifically their mean earnings. Card and Kruger (1992) also make the connection that school quantity and school quality are closely related, in that the two move together in the same direction. On the contrary, Betts (1995) is persuaded that teacher pupil ratio and teacher salary explain little of the variation in log weekly wages.

Limitations of the literature with reference to both prospects and school quality are conveyed in Long (2008). The paper goes beyond this to state that the effects of school quality can have multiple effects on economic outcomes.

In section 3, the general overview of the literature on school quality is covered through the findings of different authors presented in academic papers. Next is the section on the data which goes into detail on where the data was sourced from and the contents of the data. Some summary statistics are included in section 4 too. The outline of the empirical model lies in section 5. The model is linear in nature and its relevance to the data is made clear. One of the statistical tests which can detect conditional heteroskedasticity is the Breusch Pagan test and this is the purpose of section 6. Naturally, section 7 reports the empirical results. The empirical regressions are also displayed here. Section 8 uses the literature to support or oppose the empirical results. Section 9 draws on the limitations of the study and section 10 concludes the project.

3. Literature Review

Educational differences have long existed between school type. Understanding how these differences impact future earnings is discussed by labour economists. Given that decisions regarding human capital are made by citizens, this review is crucial. The question has consequences that include inequality of income and private costs. Policy makers look to this research when implementing policies.

The Coleman report of 1966 was the beginning of the writing on school quality. This report found that school quality played an insignificant role on student achievement as stated in Moffitt, R (1996). The controversial findings of this report have opened up various discussions regarding schooling and student achievement.

Card and Krueger have contributed greatly to this literature. State of birth is the measure of school quality in their paper. The data is split into three separate cohorts, allowing an analysis of differences in time to be conducted. When Card and Krueger (1992) observed the average rates of return to education, the figures were 5.1 percent annually for the oldest cohort and 7.4 percent annually for the youngest cohort. Card and Krueger (1992) find significant positive effects of school quality on both the average years of schooling and mean earnings of students.

The existence of omitted variables namely ability which is correlated with school quality gives rise to the endogeneity problem. This weakens the validity of a causal relationship between school quality and earnings. Without a doubt, family income and parent's education explain where a child attends school and how they view education, as referred to in Lee and Barro (2001). More income within a family generates greater access to tutoring and connections. Four propositions put forward by Card and Krueger (1998) consider the years of schooling and school quality as inputs for earnings. Card and Krueger find out that the number of years spent in education is in part determined by the school quality. An implication of this finding is that each extra year of schooling reflects greater school quality. This supports the findings of Card and Krueger (1992).

Dale and Krueger (2011) find incorporating unobserved student attributes like ability into the equation greatly diminishes the impact of school quality. The cross sectional least squares model is utililised by Dale and Krueger controlling for observable characteristics like SATs. The results show an increasing return to wages from SATs as time progresses. When the model is adjusted for unobservable student characteristics, the return becomes insignificant.

Teacher quality and class size are additional indicators of school quality. Project Star, a prominent experiment conducted in Tennessee, focuses on class size as stated in Zaharias (1999). The sample is composed of Kindergarten pupils and their teachers. Students' attainments were monitored from 1985-1989, through SATs and other cognitive tests. Over 11,000 students and teachers combined were randomly assigned into one of three groups. One group had a class size of 15 students. The other two groups had class sizes of around 22 students, one with a full-time teacher's aide. Unsurprisingly, those in the small classes performed better in tests consistently from 1985-1989 as noted in Zaharias (1999). This supports the argument that class sizes should be smaller to allow all students to flourish academically. The small class sizes were overwhelmingly advantageous to minority students from low-income households living in inner cities as expressed in Zaharias (1999). This result places attention on the inequality of income and racial discrimination that occurs within society as a whole. In summary, smaller class sizes should be analogous to greater school quality.

The long-term outcomes of Project star are considered by Chetty et al (2010). Kindergarten students with a teacher of over 10 years of experience earn $1,093 more on average at 27 years old, when compared to students with less experienced teachers. Unobserved characteristics like the ability of the teacher to engage with and motivate their students must be taken into account too. The article by Chetty et al (2010) argues that classroom environment is a key indicator of school quality and therefore test scores and earnings.

On the other hand, measures of school quality like class size do little in explaining the earnings gap according to Betts (1995). In this paper the quality of school is measured on the actual schools attended by students. Subsequently, any aggregate bias that comes from measuring school quality by state like Card and Krueger (1992) is mitigated. The link between log weekly wage and the quality of the high school is estimated. According to Betts (1995), the regressions run on each of the school quality indicators explain little of the variation in log weekly wages.

Regression discontinuity design is the method employed by Dustmann et al (2012) to pinpoint causality of school quality on future economic prospects. Students are assigned to different types of school based upon their date of birth as conveyed in Dustmann et al (2012). Their primary finding shows that the type of middle school one goes to has a small impact on secondary education. This grants students the opportunity to correct their choice of school at secondary school if need be. The paper finds value in peer effects on wages and unemployment. 80.8 percent and 30.8 percent correspond to the percentage of students who graduate middle schools of high and medium school quality respectively, as stated in Dustmann et al (2012).

Brewer et al (1996) gather panel data on students from the 1972,1980 and 1982 cohorts to identify if the effects of college quality on returns to education are constant or varying over time. This data is collected from the National Longitudinal Study of the High School Class of 1972 and High School and Beyond. A structural model is formed where college choice and subsequently college quality relies on a cost benefit analysis. As a result, college quality is an endogenous determinant of earnings. Elite private schools yield a substantial labour market premium. This premium reduces for non-elite private schools and public schools have no premium whatsoever. Increasing returns to attendance at a top school is evident based on comparing the 1972 cohort with the 1980s cohort. This finding means the effects of college quality increases with time.

To help with their review, Naylor et al. (2002) develop a statistic model on the determinants of UK graduates' occupational earnings. They discovered that earnings for males who attended an Independent school were 3 percent higher compared to those who did not ceteris paribus. For a female this figure is 3.4%. Three compelling hypotheses are brought forward to better understand the bonuses of attending an Independent school. Social networks, school quality and discrimination each have a section dedicated to explaining their relevance to earnings post university. As the degree of school quality increases so does the investment in human capital. The academic skills and discipline instilled in students at Independent schools are unmatched. This sets students up for further education. Employers actively search for students from such a background. Besides school quality, the subjects studied at school are an indicator of future earnings. Mathematics, Economics and Sciences are deemed more impressive than humanities and the classics which is expressed in Naylor et al. (2002). This paper finds variation in earnings among the class of independent schools. 113 dummy variables, one for each independent school are introduced into the earnings regression equation within Naylor et al (2002) contribution. Out of the 113 schools, 89 of them examined earnings similar to their public school counterparts. School quality cannot be approached by school type but rather on an individual basis. The study also pays close attention to the relationship linking

degree class with graduate occupational earnings. The coefficient associated with a first-class degree is 0.037 for males and 0.022 for females. This figure decreases the lower the class.

One of the more recent studies on school quality by Long (2008) makes reference to the limitations of previous studies. These limitations refer to both the outcomes of college quality and college quality itself. Long (2008) expresses that different estimation methods produce different outputs. Ordinary least squares estimator states college quality as a significant input for earnings. Nonlinear methods however, estimate insignificant effects of college quality for earnings. Another method explored is the instrumental variables method (IV). Finding instruments that satisfy both the exogeneity and relevance conditions has proven to be challenging. Throughout the literature on education, college proximity is recognized as the instrument for explaining returns to education. Nevertheless, the application of the instrument differs here. "The average quality of schools within a certain radius of the student" as stated in the paper is the instrument for the quality of the college. IV predicts highly significant results for all of the college qualities included in the regression. One of the weaknesses of this method is that the results are not valid across the board. This method can only answer for students who are less likely to travel far to attend a higher quality school. The third method from Dale and Krueger (2002) research is expanded upon. This perspective anticipates that students close in nature in terms of their goals and opportunities will have close unobservable traits. Long (2008) incorporates additional outcomes like achieving a bachelor's degree. The same approach is taken with quality through the composition of an index incorporating more quality measures. For example, the salary obtained by a professor does not explain the variation in the rate of achieving a bachelor's degree. The opposite can be said for the proportion of teacher to students. This is backed up by the reports of inconsistency in estimates between the various methods. On another note, higher tuition fees are often a signal of academic skills to potential employers. Those with higher tuition fees can be favored in the workplace.

In conclusion, the literature on school quality brings to light the lack of consensus among researchers regarding the effects of school quality on future outcomes. Ability bias is one of the challenges faced when estimating the link between school quality and outcomes like earnings. Moreover, a great deal of uncertainty about this topic still exists. Lastly, the effects of school quality have multiple effects on economic outcomes.

4. Data Description

DATA

The dataset utilised here is from the Project Star (Student Teacher Achievement Ratio) Experiment. Project Star is a large-scale randomised experiment conducted in Tennessee between 1985-1989. The type of data is a panel dataset. The experiment was funded by Tennessee legislators and the Tennessee Department of Education. The data comes from the Harvard Database. All the information in this section can be found online at (Tennessee's Student Teacher Achievement Ratio (STAR) project, 2008). The project consisted of 11,601 students and teachers combined to demonstrate how the size of a class can greatly contribute to the learning outcomes of the students. The school children ranged from Kindergarten up to grade 3 who attended schools in Tennessee. The students were randomly assigned into one of three groups. The first group was the small class size of about 15 students and this is the treatment group. The second group is the regular class with around 22 students and this is the control group. The final group is the regular class size of about 22 students plus a full-time teacher's aide. 79 schools participated in this study. The teachers were randomly assigned into one of the three class types too. The assessment of the students included achievement in tests on an annual basis. Student achievement was measured using the Stanford Achievement Test (SAT) for reading, mathematics, listening and word skills. Basic Skills First tests were introduced in first grade and college entrance examinations, ACTs and SATs, were taken in high school. The results of this experiment showed that students in smaller class sizes performed better in tests.

There are four data files as noted by Finn et al (2007). One of the files is on 21 schools comparable to STAR schools which did not participate in Project Star. The file on the STAR students is the focus of this project. This data consists of 11,601 observations and 379 variables. The third file is on the high schools participating in the STAR project and the fourth is on the Kindergarten to grade 3 STAR schools. Due to there being an overwhelmingly large number of variables, only some of the variables were selected for this project.

VARIABLES

Whilst the outcome variables which will be discussed in this project are SAT scores in maths and reading for all students through K-3, data on study skills and listening scores is also provided in the dataset. Additionally, ACT and SAT scores will also be indicators for economic outcomes. It is important to note that many of the variables have missing data, creating potential bias in the results.

The main independent variable of interest is class size. Other variables included in the data are race and gender pertaining to both students and teachers. Classroom type for kindergarten year group, grade 1, grade 2 and grade 3 are reported separately. The variable small is a dummy variable indicating students who are assigned to a small class. Likewise, regular and regular plus aide dummy variables have been generated. Black is a dummy variable denoting the student's race is black if the variable is equal to 1 and another race if equal to zero. Furthermore, years of total teaching experience in kindergarten is treated as a teacher control in the regressions. The years of teaching experience can be used as a proxy for teacher quality. These are two variables that relate to the teachers in the empirical results section. One is a categorical variable on the years of teaching experience which ranges from 0-27 years with no entries for 23, 25 and 26 years. The gender of the teacher is the other variable referring to the STAR teachers.

Lastly, the variable free lunch is useful for indicating family background and thus is a part of the regressions.

With regard to demographic variables in relation to the schools, these offer such insight in how where you attend school determines the performance of the student to some degree. This is particularly helpful in explaining within state differences, being that all the schools in Project Star are in Tennessee. The demographic variables consist of inner city, rural, suburban and urban.

SUMMARY STATISTICS

	Observations	Mean	Standard deviation	Min	Max
Gender	11,587	-	-	-	-
Race	11,467	-	-	-	-
SAT Maths score	489	519.3252	105.7785	200	800
Total SAT (Verbal And Maths score)	489	1047.096	197.8402	400	1560
Total ACT (English, Reading, Maths and Science score)	3,754	76.02184	17.966	35	136

Class type

	Kindergarten	Grade 1	Grade 2	Grade 3
Small:	1900	1925	2016	2174
Regular:	2194	2584	2329	2085
Regular with aide	2,231	2320	2495	2543

Kindergarten was not compulsory in Tennessee, so some students joined the STAR experiment in grade 1 rather than Kindergarten which explains why the number of students in grade 1 is higher than in Kindergarten.

5.Empirical Model

The model is derived from Krueger (1999). The model is as follows:

$$Y_{ij} = aS_{ij} + bF_{ij} + \varepsilon_{ij}$$

where Y_{ij} is the achievement level of student i who attends school j. S_{ij} is a vector of school characteristics and F_{ij} is a vector denoting the family background of the student i. ε_{ij} is the error term. a and b are the associated slope coefficients for each of the independent variables. The specification is linear and will be interpreted as such.

In the empirical regressions, there is limited data on the family background of the student (F_{ij}) but the dataset includes a variable named free lunch which provides an idea of the income level of the family. With respect to S_{ij}, for grades K-3 at least, this consists of class size and demographic of the school. In addition to this, the gender and race of the teachers and student are present in the regressions. Furthermore, teacher quality in the form of experience and degree level are contained in S_{ij}. Y_{ij} encompasses the academic achievements of students through their scores in tests taken annually during the project. SAT and ACT marks are also outcome variables of interest.

6. Econometric Test

In order to ensure the OLS estimator is unbiased and consistent, the classical linear assumptions must be satisfied. Assumption 1 holds which states that the model is linear in parameters. Given the nature of Project Star, the sampling is random satisfying the second assumption. Based on the fact that both students and teachers are randomly assigned into their classes, the self-selection bias is not a cause of concern.

There was some doubt when it comes to assumption 5 which is conditional homoskedasticity. Upon carrying out the Breusch Pagan test, it was evident that heteroskedasticity was present in the model. Therefore, we reject the null hypothesis of conditional homoskedasticity and find evidence to suggest that the variance of the error is heteroskedastic. In this case, the variance of the error term is dependent on the value of the explanatory variables. Consequently, we use the heteroskedasticity robust standard errors. The failure of conditional homoskedasticity has consequences. It causes the standard errors to be incorrect and produces invalid t and F statistics. Moreover, the OLS estimator is no longer the best linear unbiased estimator. Consequently, the heteroskedasticity robust standard errors are reported.

7. Empirical Results

Table 1: Ordinary Least Squares regression for the Kindergarten cohort with a small class size.

Dependent variable: Maths score in Kindergarten.

	Small	t	Regular	t	aide	t
	8.78	5.94	-3.50	-2.51	-4.79	-3.55
	(1.48)		(1.40)		(1.35)	
Black	-8.44	-3.83	-8.54	-3.87	-8.46	-3.83
	(2.20)		(2.21)		(2.21)	
Female	7.06	5.36	7.08	5.36	6.96	5.27
	(1.31)		(1.32)		(1.32)	
Teacher experience	0.49	4.11	0.46	3.82	0.50	4.23
	(0.12)		(0.12)		(0.11)	
Black teacher	8.12	3.85	7.96	3.77	7.51	3.54
	(2.10)		(2.11)		(2.12)	
Inner city	-6.10	-2.49	-6.41	-2.62	-6.36	-2.60
	(2.44)		(2.45)		(2.44)	
Rural	-4.38	-2.73	-4.55	-2.82	-4.99	-3.10
	(1.61)		(1.62)		(1.61)	
Free lunch	-16.89	-11.30	-16.95	-11.31	-16.84	-11.24
	(1.49)		(1.50)		(1.50)	

Motivation	0.01	0.05	0.03	0.11	0.02	0.08
	(0.27)		(0.28)		(0.28)	
Observations: 4756						
R squared	0.078		0.0719		0.073	

The figures above those in parentheses are the coefficients. The figures in parentheses for all tables in this paper are heteroskedasticity robust standard errors. Where applicable, numbers are rounded to 2 decimal places. Wherever teaching experience is mentioned, it denotes years of teaching experience.

Table 2: Ordinary Least Squares regression for grades 1-3 with a small class size.

Dependent variable: Maths score for grade 1,2 and 3.

Grade	Coefficients	Robust standard errors	t ratio
1	10.25	(1.20)	8.52
2	7.38	(1.54)	5.90
3	6.15	(1.07)	5.72

The explanatory variables are identical to Table 1.

For all the students in kindergarten to grade 3 for which there is available data, being in a small class has a positive effect on maths score and when this is compared to both types of regular classes, it is evident how a small class size greatly increases the math test score. This result should make policy makers think about reducing the number of students in any class. The extent to which this can occur is dependent on the budgets for individual schools as well as the resources at their disposal. The same trend applies to reading test scores. Nonetheless, the R squared in some of the cases is small even after controlling for teacher characteristics, student characteristics and the neighbourhood in which the school is at. This could indicate measurement errors exist.

Through conducting a t test for significance at the 5% level of significance, it can be concluded that the coefficient for small class type is significant and as such there is sufficient evidence to suggest that the variable small is statistically significant. The years of teaching experience the teacher has does not appear to explain much of the variation in the maths test score. Betts (1995) also finds that teacher effects are insignificant with respect to wages. Precisely, the proportion of teachers with master's degree or above and the ratio of teachers to students are not strong determinants of future wages of the students.

Contrastingly, Chetty et al (2014) reports that teachers which exhibit high value-added characteristics, have a positive but also important impact on students well into their adulthood. The variable free lunch is a good indicator for the economic status of the household and as expected the coefficient is negative. This is due to the fact that typically students who receive free lunch come from a low-income family. This is interpreted as students who receive a free lunch at school obtain a lower test score than students who do not receive free lunch. Family background and income

therefore plays a critical role in a child's education as stated in Lee and Barro (2001). Demographic variables alongside race, gender and entitlement to free lunch cause more variation in the maths test score than other variables like motivation. The variable motivation relates to how the students feel.

Table 3: Ordinary Least Squares regression of Reading Test Scores on small class types.

Variables	Kindergarten	Grade 1	Grade 2	Grade 3
Small	5.59(0.96)	10.61(1.53)	7.85(1.25)	7.22(1.04)
Black	-7.38 (1.42)	-16.12 (2.05)	-16.67(1.77)	-13.00(1.54)
Female	5.58 (0.87)	11.36 (1.35)	9.56(1.13)	7.73(0.98)
Teacher experience	0.45 (0.08)	0.33(0.08)	0.10(0.06)	0.13(0.06)
Black teacher	1.76 (1.26)	-1.79(2.01)	3.28(1.44)	2.33(1.42)
Inner city	0.64 (1.55)	-10.99(2.20)	-6.14(1.71)	-1.17(1.71)
Rural	-4.85 (1.13)	2.06(1.65)	3.43(1.44)	1.00(1.18)
Free lunch	-13.24(0.95)	-29.23(1.55)	-21.22(1.35)	-17.49(1.11)
Motivation	0.13 (0.18)	0.86(0.17)	0.31(0.15)	0.06(0.13)

For the variable black teacher, this is white teacher for grades 1, 2 and 3. The number of observations for each grade beginning with kindergarten is as follows: 4693, 5557, 5582 and 5513.

Table 4: Ordinary Least Squares regression with grade 4 Maths Stanford Achievement test (SAT) as the dependent variable.

	coefficient	t ratio
Small	5.68(1.42)	4.01
Black	-14.06(1.91)	-7.38
Female	8.26(8.26)	6.13
Teacher experience	0.02(0.08)	0.21
Free lunch	-22.18(1.55)	-14.33

All the explanatory variables in this Table are concerning the grade 3 cohort.

Table 5: Two Ordinary Least Square Regression where grade 4 Maths SAT is the dependent variable for both.

	Regular		Regular with aide	
	coefficient	t	coefficient	t
Small	-1.05(1.46)	-0.71	-4.50(1.42)	-3.17
Black	-14.02(1.91)	-7.34	-14.051.91)	-7.37
Female	8.31 (1.35)	6.16	8.28(1.35)	6.15
Teacher experience	-0.002(0.08)	-0.03	0.004(0.08)	0.05
Free lunch	-22.27(1.54)	-14.38	-22.14(1.55)	-14.30

Number of observations: 4293.

R squared: 0.0946, 0.0967. The first R squared corresponds to the regression on the left-hand side of Table 5 and the second corresponds to the second regression in Table 5.

Students in grade 4 achieve higher marks in maths if they previously attended a small class in grade 3 during the Project Star experiment. Since all students return to regular classes in grade 4, the effects of a small class size in grade 3 are seen in grade 4 maths score. Both types of regular classes have coefficients which are negative, meaning those students obtain lower scores than those in small class. Students in regular classes with a teacher's aide perform considerably worse than those without an aide. It can be concluded that having a teacher's aide in the classroom has no benefit at all for fourth graders taking a maths test.

To demonstrate the relationship between SAT maths scores in high school and the maths score in grade four, the correlation of the two variables is 0.6328 which is moderately high. This can be used to support the notion that the academic achievements of an individual in their early childhood can be used to predict future achievements. Although, Dustmann et al (2012) makes the valid point that so long as one starts to receive higher quality schooling by secondary school, the positive effects of school quality will be experienced through higher earnings.

Table 6: Ordinary Least Squares Regression with ACT as the dependent variable.

	Kindergarten	Grade 1	Grade 2	Grade 3
Small	0.09(0.71)	0.21(0.67)	0.32(0.66)	0.78(0.65)
Black	-12.39(1.07)	-11.44(0.997)	-12.45(0.99)	-13.03(0.998)
Female	-0.93(0.69)	-0.86(0.66)	-0.43(0.64)	-0.46(0.64)
Inner City	-0.09(1.26)	1.13(1.26)	0.93(1.18)	0.35(1.21)
Suburb	3.30(0.91)	3.09(0.85)	2.25(0.86)	2.73(0.86)
White Teacher	-0.43(1.06)	2.04(0.996)	1.27(0.87)	0.79(0.96)
Teacher experience	0.03(0.06)	0.06(0.04)	0.044(0.357)	0.09(0.04)
Free lunch	-4.39(0.80)	-5.73(0.77)	-6.39(0.75)	-5.43(0.76)
Observations	2416	2661	2726	2754
R squared	0.1493	0.1515	0.1730	0.1697

Table 7: Ordinary Least Squares Regression with ACT as the dependent variable.

			R squared
Regular	-0.15(0.68)	-0.22	0.1693
Regular with aide	-0.63(0.65)	-0.96	0.1695

Observations: 2754.

For Kindergarten and grades 1, 2 and 3 the coefficients on small are positive. This can be interpreted as an individual in a small class improves the ACT score obtained in high school. Hence, the effect of a small class on total ACT score is positive. Through careful observation, it can be seen that the coefficient on small increases as the grade increases. In Kindergarten the figure is 0.09 and by grade 3 this figure is 0.78. Also, students who attend schools in the suburbs achieve better test scores than those whose schools are in the inner city.

The regressors reported in Table 6 are the same regressors for Table 7. For simplicity I have only included the coefficients for the two regular class types. From the table it is evident that regular classes give students a higher ACT score when compared to regular plus aide classes. This demonstrates that the presence of a teacher's aide actually puts the students at a disadvantage for the ACT test. Students in regular and small classes outperform those in regular classes with a

teacher's aide. This result places emphasis on the classroom environment and how students can feel more confident when the class is smaller. Nonetheless, in terms of magnitude and size, the coefficients are small.

Table 8: Ordinary Least Squares regression for black females in grade 3.

Dependent Variable: High School SAT maths score

	coefficient	t ratio
Small	7.27(8.68)	0.84
Grade 8 maths score	1.66(0.11)	14.89
Black	-37.43(10.71)	-3.50
Female	-16.25(9.39)	-1.73
Urban	-15.87(17.30)	-0.92
Rural	-7.55(12.09)	-0.63
Teacher experience	-0.45(0.598)	-0.76
Free lunch	-14.64(15.27)	-0.96

Observations:225

R squared:0.6199

Table 9: Ordinary Least Squares regression for white females in grade 3.

Dependent Variable: High School SAT maths score

Small	5.598(8.77)	0.64
Grade 8 maths score	1.71(0.11)	15.13
White	21.42 (11.06)	1.94
Female	-18.85(9.45)	-1.99
Urban	-12.91(17.50)	--0.74
Rural	-4.71(12.24)	-0.38
Teacher experience	-0.47(0.61)	-0.77
Free lunch	-23.19(15.18)	-1.53

Observations:225

R squared:0.6102

For a white female, the coefficient on small is 5.598. The coefficient on white dummy variable is 21.42, which is a contrast to the coefficient on black of -37.43. Race plays a large role in the test score achieved by students. Based on this data, whites outperform blacks in normal class settings even after gender, teacher characteristics and demographics have been controlled for. This brings racial inequality of education to the forefront. Policy makers should make this a priority. But in small classes, blacks have higher improvement rates in test scores than whites. This shows the effectiveness of small classes. Out of all the regressors in this regression, only grade 8 maths score and the race dummy variables are significant at the 5% level of significance.

8. Comments

Fourth grade teachers filled out a questionnaire relating to student participation. One of the questions is with regard to paying attention. With regard to the fourth-grade cohort, approximately 33% of students assigned to small classes in kindergarten are well focused in the classroom. This figure reduces to around 27% for both types of regular classes. By grade 4 all classes returned to their normal class size. Despite this, the small class effects on effort and initiative of students can still be perceived as reported by Finn et al (1989).

The dependent variable selected by Dearden et al (2002) is log hourly wage rates. Estimates for females and males are examined separately, at age 23 and then again at 33 years old. Having controlled for family background and ability, pupil teacher ratio is insignificant for males at both time period. While ability and school type are initially insignificant at age 23, they become significant at age 33 according to Dearden et al (2002). This highlights that the effects of school quality measures differ depending on the age. From Dearden et al (2002) findings, it is noted that test scores obtained at 11 years old are significant for a female aged 23 when family background is controlled for. A key finding is that a 1% decrease in the pupil teacher ratio increases the wages of females by 1% at age 33 according to Dearden et al (2002). This stands out especially as a great proportion of the literature looks at earnings for students who have recently graduated.

This implies that the effects of school quality indicators may not always be felt immediately but rather in the long run. Card and Krueger (1996) also take the stance that the determinants of labour market performance like school quality are only uncovered with experience. A suggestion would be for the school quality effect on earnings to be assessed on individuals who have been in the labour market for 20 and 30 years and so forth.

This aligns with Krueger (1999) where it states measured teacher characteristics within Project Star explain a small portion of student attainment.

Chetty et al (2011) links the Project Star experiment with administrative records to assess the long-term impacts of class size. Chetty et al (2011) take the analysis of variance approach to measure the spread of earnings of the varying class sizes for the kindergarten cohort. The results of this approach demonstrate that kindergarten classes have significant effects on earnings. Whether higher test scores impact earnings or not is unclear. As noted in the empirical analysis section, the impacts of class size on test scores lessen as time passes. This observation is supported by Krueger and

Whitmore (2001) who state that by grade 8 class size effects are statistically insignificant. In an attempt to explain why test scores, become more important in adulthood earnings despite their insignificance in high school, Chetty et al (2011) considers the part non-cognitive skills play. Through experimenting, a high correlation between earnings and non-cognitive measures like initiative and engaging behaviour is found.

James et al (1989) perceives choice of major and grade point average as contributing massively to the variation in earnings. A stronger statement is that these two variables cause a wider variation in earnings than family background. Put another way, choice of major can overcome the setbacks of unfavourable family background.

Drop-out rate is to some degree a short-term consequence of school quality. The probability of students dropping out of university is estimated by Smith and Naylor (2001) using a probit model. There are reasons besides school quality that cause individuals to drop out. Having said this, one of the variables looked at is teaching quality of universities. The marginal effect this variable has on drop out rates are statistically significant and negative. As teaching quality improves, the likelihood of students dropping out lessens as stated in Smith and Naylor (2001). This is evidence supporting teaching quality matters for outcomes, both positive or negative. In this case drop out rate is a negative outcome. Hence, attention should be given by schools on improving the learning within classrooms.

School type is also included in the model and shockingly the coefficient on independent schools is larger than that of Local Education Agencies (LEA) schools. While dropping out of school is a short-term issue, the unemployment rate is a long term one. According to Smith and Naylor (2001), a five-percentage point increase in the local county level of an individual will lead to a one percentage point increase in the drop out rate. This stands out because differently from the rest of the literature, the direction of causality goes from the outcome to factor. More importantly, it demonstrates that school quality is not exogenous and some of the effects of school quality are pre-determined.

9. Limitations of the Study

Omitted variable bias is clearly present, with variables such as ability not included in the list of variables as it is unobserved and hence difficult to measure. This is a problem discussed in most of the literature. Endogeneity is another problem that weakens the validity of the result. Ideally, an instrument variable would be used as an exogenous variable to replace the endogenous variable. Finding instrumental variables that satisfy both the relevancy and exogeneity conditions is difficult to achieve in labour economics as mentioned in Long (2008). This project would be better if data on variables such as IQ and whether the student has extra tutoring sessions were included. Gathering data from parents on whether their children take tutoring classes requires compliance on their part. This could potentially pose as a challenge. The data used by Chetty et al (2011) examining the long-term effects of Project Star could not be accessed. For this reason, regressions could not be carried out as planned. This would have strengthened the hypothesis that small class sizes are important for determining future economic prospects beyond schooling age.

A second limitation is to do with the endogeneity of school quality. In the error term lies omitted variables such as the education of the parents which are valuable in explaining the educational outcomes of their children.

10. Conclusion

To conclude, there is no clear consensus among economists. As a result, further research is needed on the research question. To answer the question pertaining to the effect of school quality on future prospects, the best response is to say uncertain.

Most of the literature concerning this question is based on schools in either the United States or the United Kingdom. It would be advantageous for more countries to do their own research so comparisons between countries can be made. In addition, it could help in gaining a better understanding of the topic, if patterns are observed for example.

The different specifications used among economists with regard to the school quality literature will affect the findings. There is great ambiguity in both the effect of school quality on future outcomes and whether the proxies used for school quality are appropriate. Previously, school quality was primarily viewed as expenditure per pupil. With time, the literature has considered whether this is truly related to future economic prospects.

School quality can mean so many things ranging from school resources to type of school. Regardless of this, as shown by Naylor et al (2002), there is variation in earnings among independent schools. Moreover, outcomes for males and females differ, even when all other factors are identical. Blacks and whites have different experiences when it comes to the effect of school quality. It is evident both racial and gender discrimination start from the classroom.

In many cases in the literature as well as this project the “quality” of a school has been simplified to allow models to be specified. It remains that the effects of quantity of schooling are easier to study and interpret.

Whilst the majority of the literature associates school quality with cognitive skills, for example, mathematics which can be assessed and scored, the effect of non-cognitive skills on future economic outcomes remains untouched. Scoring well in tests and examinations are a signal to employers of higher ability but this measure on its own does not explain school quality in its entirety. In the labour market, companies are placing greater focus on the skills and attitudes of job seekers instead of solely on academic achievement. Unsurprisingly, non-cognitive skills are more subtle than cognitive so it is often overlooked or misunderstood. It is therefore possible that school quality has a greater effect than is being portrayed in the literature. It may be time to expand the ways in which school quality is measured. The “quality” of a school in which one attends is certainly key in determining their future economic prospects. Finally, the effects of school quality are many.

11. References

American Federation of Teachers (1999) *Project STAR, The Story of the Tennessee Class- size Study.* Available at: Project STAR: The Story of the Tennessee Class-size Study by Jayne Boyd-Zaharias - American Educator, Summer 1999 (aft.org) (Accessed: November 2020).

Betts, J. (1995) 'Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth', *The Review of Economics and Statistics*, 77(2), pp.231-250.

Brewer, D., Eide, E. and Ehrenberg, R (1996) 'Does It Pay to Attend an Elite Private College? Cross Cohort Evidence on the Effects of College Quality on Earnings ', National Bureau of Economic Research, pp.1-23. doi:10.3386/w5613

Card, D. and Krueger, A (1992) 'Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States ', Journal of Political Economy, 100(1), pp.1-40.

Centre for Research and Analysis Migration (2012) *The Long term effects of School Quality on Market Outcomes and Educational Attainment*. Available at: results_February_2012_corrected.xls (cream-migration.org) (Accessed: 2 December 2020).

Chetty, R. et al (2011) 'How Does your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star', The Quarterly Journal of Economics,126(4) ,1593-1660. doi:10.1093/qje/qjr041.

Chetty, R. et al. (2014) 'Measuring the Impacts of Teachers II: Teacher Value- Added and Student Outcomes in Adulthood', American Economic Review,104(9), pp.2633-79.

Dale, S. and Krueger, A. (2002) 'Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables', The Quarterly Journal of Economics, 117(4), pp. 1491-1527. doi:10.1162/003355302320935089.

Dale, S. and Krueger A (2011) *National Bureau of Economic Research*. Available at: (Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data | NBER /Accessed: October 2020).

Dearden, L. et al. (2002) 'The Effect of School Quality on Educational Attainment and Wages', The Review of Economics and Statistics, 84(1), pp.1-20.

Finn, J. et al. (1989) ' Carry-over effects of small classes', Peabody Journal of Education, 67(1), pp.75-84.

Finn, J. et al. (2007) *User manual | Project STAR and Beyond: Database User's Guide*. Available at: Project STAR and Beyond: Database User's Guide | Manualzz (Accessed: 15 January 2021).

James, E. et al. (1989) 'College Quality and Future Earnings: Where Should You Send Your Child to College?', The American Economic Review, 79(2), pp.247-252).

Krueger, A. (1999) 'Experimental Estimates of Education Production Functions', The Quarterly Journal of Economics, 114(2), pp.497-532.

Krueger, A. and Whitmore, D (2001) 'The Effect of Attending a Small Class in the Early Grades on College- Test Taking and Middle School Test Results: Evidence from Project STAR', The Economic Journal, 111(468), pp.1-28

Lee, J. and Barro, R. (2001) 'Schooling Quality in a Cross- Section of Countries', Economica, 68(272), pp.465-488

Long, M. (2008) 'College Quality and early adult outcomes, Economics of Education Review', 27(5), pp.588-602.

Moffitt, R. (1996) 'Symposium on School Quality and Educational Outcomes: Introduction', *The Review of Economics and Statistics*, 78(4), pp.559-561. doi:10.2307/2109947.

National Bureau of Economic Research (1996) *Labor Market Effects of School Quality: Theory and Evidence.* Available at: Labor Market Effects of School Quality: Theory and Evidence | NBER (18 Feb 2021).

Naylor, R., Smith, J. and McKnight, A. (2002) ' Why is There a Graduates Earnings Premium for Students from Independent Schools', *Bulletin of Economic Research, 54(4)*, pp. 315-339.

Smith, J. and Naylor, R (2001) 'Determinants of Degree Performance in UK Universities: A Statistical Analysis of the 1993 Student Cohort', *Oxford Bulletin of Economics and Statistics*, 63(1), pp.29-60.

Tennessee's Student Teacher Achievement Ratio (STAR) project (2008) Available at: Tennessee's Student Teacher Achievement Ratio (STAR) project - Project Star Dataverse (harvard.edu) (Accessed: 15 January 2021).

The ANNALS of the American Academy of Political and Social Science (1998) *School Resources and Student Outcomes*. Available at: School Resources and Student Outcomes - DAVID CARD, ALAN B. KRUEGER, 1998 (sagepub.com) (Accessed: 10 November 2020).