Introduction

Context

Government spending on US education has been a long standing debate. In 1966, James S. Coleman conducted a survey for the federal government to address a section in the Civil Rights Act of 1964 "concerning the lack of availability of equal educational opportunities for individuals by reason of race, color, religion, or national origin." After collecting data on 650,000 students and teachers, he wrote a 700-page report "Equality of Educational Opportunity" stating that school resources matter less than family in influencing student outcome. Many took this as a sign that additional spending on education would make little difference. However, Coleman also found evidence of the "achievement gap," which illuminates test score disparities between rich and poor students [5]. It seems that if the money is spent in the right way, for example, in bridging the gap between school quality, extra spending does make a difference [1].

Literature

The Coleman report continues to spur further studies on school spending and student success. A publication from 2015 "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms" found that a "10 percent increase in per-pupil spending... leads to 0.27 more completed years of education, 7.25 percent higher wages, and a 3.67 percentage-point reduction in the annual incidence of adult poverty." In addition, these effects were even greater for students from low socioeconomic backgrounds [3]. Another study "School Finance Reform and the Distribution of Student Achievement" from 2018 draws on student-level data to identify the effects of school finance reforms that began in 1990 on "relative achievement of students in high- and low-income school districts." They found that school finance reforms on spending in low-income school districts matter as well, as they lead to improvement in student achievement, measured by test scores [2].

Research Questions

For our analysis, we investigate whether familial or educational factors are larger determinants of student success.

First, we define familial and educational factors by categorizing all predictors of interest under these two groups. Using the familial predictors, we can evaluate whether there is credit to Coleman's claim that family is a larger determinant than school resources of student success. For the educational predictors, we can analyze the association of school finances and student success. However, we will not be looking directly at school expenditure and revenue since the data is not available in the dataset used. Instead, we will analyze variables that could be correlated with expenditure, such as building conditions, teacher salary, and access to resources. Since public school revenue is related to local taxes, we will also be looking at socioeconomic factors on the school level.

Next, to measure student success, we chose to evaluate more than just high school standardized test scores, since we believe that student success is not something that can be measured by a single test while a student is still developing. By analyzing variables that are determined years after high school, we hope to get a more holistic sense of student outcome. Therefore, we measure student success through standardized test scores in 10th grade, level of education attained 10 years after 10th grade, and socioeconomic status 9 years after 10th grade.

Data

The data come from National Center for Education Statistics (NCES) from the Education Longitudinal Study of 2002. In this study, students were surveyed three times: in 2002, as high school sophomores; in 2006,

two years after graduation; in 2012, eight years after graduation. The data is free for public use, with the agreement of the NCES Data Usage Agreement. There are over 16,000 observations on the student level [4].

Variables and Exploratory Data Analysis

Because of the number of variables used in the models, there are many observations that have missing data for at least one of the variables. Around half of the variables were missing less than 1000 observations, while some were missing around 3000-4000. The variable with the highest amount of data missing is the variable for how far a math/English teacher pushed a student to go in school, with around 4000 missing observations for both math and English teachers. e use the `mice` package to impute the data using the predictive mean-matching method. To support our results with imputed data, we perform sensitivity analysis with non-imputed data, which still showed similar significance in the three models for the predictors of how far a math/English teacher pushed a student to go in school.

Due to the small sample sizes within the levels of combinations of the categorical variables and to better take into account their ordinal nature, we transform them into numeric variables. The variables of how far the parents pushed the students to go in school and the parents' highest level of education are transformed into numeric variables that represent the number of additional years of education past 9th grade for each level indicated. We acknowledge the limitations of this, as there are assumptions made that may not be accurate for all observations. For example, for levels of education that denote only attending, and not completing, high school or college, we arbitrarily set a number of years of education that would fall under the category, as the exact number of years is unknown.

Additionally, there is a set of categorical variables that denote how much learning is hindered by a particular environmental factor from a scale of 1-4, where 1 denotes no learning hindrance. Again, sample size is a problem, as our models include many categorical variables. Therefore, to optimize the number of levels that have at least 1 observation and since we do not assume linearity between each level of the variable, we collapse these categorical variables into binary indicators, instead of converting them to numeric variables. We encode each learning hindrance variable into a binary of whether or not students face any learning hindrance (levels 2-4) from the specific environmental factor. Although this results in some loss of information that's gained from the differences between levels 2, 3, and 4, we believe that this is necessary to guarantee sufficient sample sizes to generate the models.

From Figure 1, we see that for higher levels of educational attainment, the proportion of two-parent families increases, while the proportion of those with a single parent or absent parent decreases. This shows an association between family factors and student success, perhaps somewhat backing part of Coleman's claim. In the following section, we will control for confounding variables and analyze more thoroughly the importance family factors, such as family composition, in comparison to educational ones.

On the educational side, previous literature mentions how increased spending to decrease class sizes helps to bridge the achievement gap between wealthy and poor districts. To represent the wealth of a district, we are using the percentage of students that receive free lunch, since we assume it measures general socioeconomic status of the area of the school.

Figure 2 shows an increased association between lack of space as a learning hindrance and student test scores for schools with more than 75% of students with free lunch, whereas the association seems positive or very small in schools with lower percentages of students with free lunch. This might suggest a magnified correlation between these two variables in poorer districts, aligning with previous findings. Therefore, we will include an interaction between free lunch percentage and whether learning is hindered by a lack of space in our models.

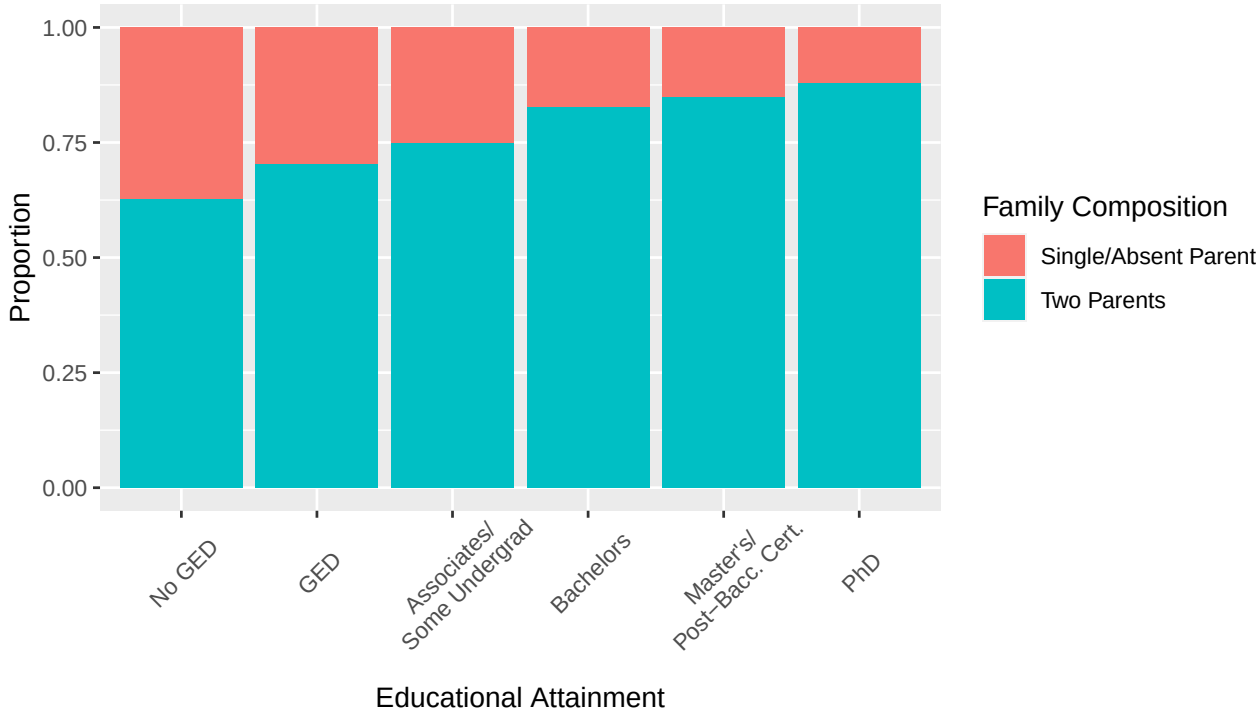

Figure 1: Familial Predictors: Positive Relationship between Two Parent Families and Higher Education Attainment

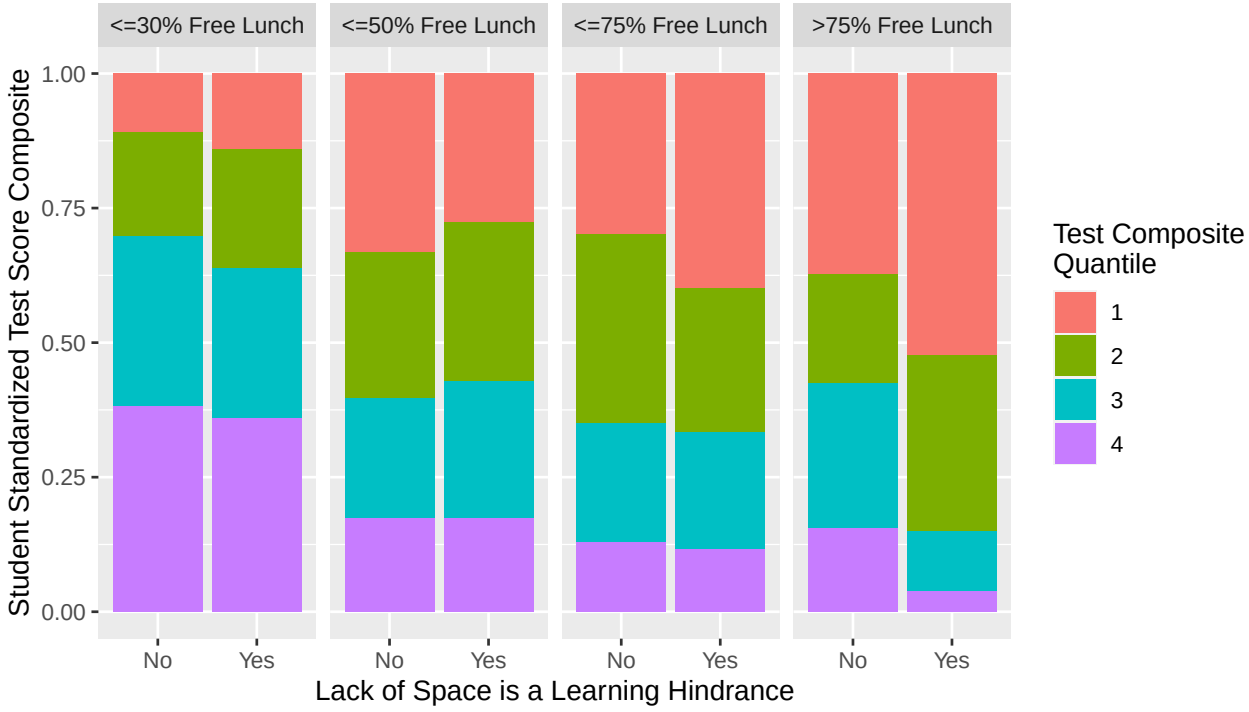

Figure 2: Educational Predictors: Lack of Space Negatively Correlated with Test Scores, Especially as School Wealth Decreases

Sources

[1] https://www.npr.org/sections/ed/2016/04/25/468157856/can-more-money-fix-americas-schools

[2] https://equitablegrowth.org/can-school-finance-reforms-improve-student-achievement/

[3] https://www.nber.org/papers/w20847

[4] https://fordhaminstitute.org/national/commentary/education-longitudinal-study-2002

[5] https://hub.jhu.edu/magazine/2016/winter/coleman-report-public-Education/