The Effect of School Funding Inequality on Student Achievement Using Causal Inference Methods

Elijah Tamarchenko '23

1. Introduction

Income inequality in the U.S is currently at an all time high and actively increasing, which in turn is causing an increase in the inequality in public school funding (Boustan et al., 2013). In the 1960's Coleman (1968) published the Coleman report, a massive research paper which found that factors outside of student body composition - factors such as school funding, resources, teacher quality - had very little to do with student achievement. Since then, researchers have had mixed results in trying to disprove most of the results from the Coleman report, and in attempting to prove that school related factors have a heavy influence on the achievement of their students (Downey and Condron, 2016). Neymotin (2010) found that there was little evidence that changes in school funding had any effect on the graduation rates of students in public high schools in Kansas, and Han et al. (2021) found that even though increased funding had an effect on student achievement, this improvement was only due to improvement of the already high achievers, and had little impact on the students who were struggling. In contrast, Kreisman and Steinberg (2019) found that in Texas, an increase in funding had a large effect on graduation rates and college enrollment, and that this change was actually more prominent in poorer districts where more struggling students are present. This finding was supported by Lafortune et al. (2018) who found that increasing funding in low-income districts improved student performance in the long run. Thus the problem of the impact of funding on student achievement is still a debated one, with many results on either side of the argument, although recent research tends to trend towards affirming the relationship. The problem of determining whether school funding has an effect on the achievement of school is an essential one to tackle, especially since there have been recent trends towards prioritizing school funding towards schools with higher student achievement, which could potentially widen the student achievement gap and increase segregation and inequity in the U.S educational system (Ostrander, 2015). I agree with the criticisms of the Coleman report, and hypothesize that a below average amount of funding in a school is a direct cause of lower student proficiency, specifically as measured by the MCAS (Massachusetts Comprehensive Assessment System). I create a causal graph using previous literature and expert knowledge, and then use backdoor adjustment and AIPW (Augmented Inverse Probability Weighting) in order to test this causal relationship, and find that decreases in school funding are indeed a direct cause of decreases in the student proficiency in the school.

2. Preliminaries

An ADMG (Acyclic Directed Mixed Graph) G = (V,E) is defined as a graph that contains only directed and bidirected edges, with at most one directed and one bidirected edge in between each pair of vertices, and no directed cycles in G, i.e. such that there is no sequence $V_i \rightarrow V_k \rightarrow \cdots \rightarrow V_i$ in the graph. A directed edge from $V_i \rightarrow V_j$ in this case will be inferred as stating that V_i is a potential directed cause of V_j, and likewise a bidirected edge from $V_i \leftrightarrow V_j$ implies that both V_i and V_j have a potential shared direct cause U_k, which is unmeasured and thus omitted from the graph. According to these specifications of the ADMG, the joint distribution $p(V)$ factorizes with respect to G according to the following distribution function:

$$p(V) = \prod_{i=1}^{k} p(V_i | \text{pa}_{\mathcal{G}}(V_i))$$

where $\text{pa}_{G}(V_i)$ are the parents of the vertex V_i in G, defined as having a directed edge into V_i. In simple terms, this function for $p(V)$ states that the joint distribution is equal to the multiplicative sum of the conditional probability of each vertex in the graph given their parents. Thus the joint distribution is characterized by the average causal effect the parents of vertices have on their children.

The conditional independence statements implied by G can be found using m-separation (a measure of disconnectedness in graphs) using the ordinary global Markov property, which states that for all disjoint subsets of vertices X,Y,Z in G,

$$(X \perp\!\!\!\perp Y \mid Z)_{\text{m-sep in G}} \Rightarrow X \perp\!\!\!\perp Y \mid Z \text{ in } p(V)$$

where $X \perp\!\!\!\perp Y \mid Z$ implies that X is conditionally independent of Y given Z. This property thus states that conditional independence statements that hold in G through mseperation will also hold in the joint distribution $p(V)$ that we defined earlier. We can further refine this equation through the faithfulness assumption, which states that data coming from a system which adheres to a causal graph will contain the same relations which we have discovered using our m-separation algorithm. Thus,

$$(X \perp\!\!\!\perp Y \mid Z)_{\text{m-sep in G}} \Leftrightarrow X \perp\!\!\!\perp Y \mid Z \text{ in } p(V)$$

meaning that an independence relation in G implies an independence relation in $p(V)$ and vice-versa.

The Average Causal Effect (ACE) is a measure of the total effect of a treatment A on outcome Y, along all of the causal pathways, defined by the function

$$ACE = \mathrm{E}[Y(a)] - \mathrm{E}[Y(a')]$$

where a denotes the presence of a treatment and a' denotes the absence of the treatment. The ACE can be interpreted as the amount with which we expect Y to differ between the presence of the treatment and the counterfactual (the absence of treatment).

3. Methods

3.1 Data

In order to estimate the causal effect of school funding on student achievement, we use data from Massachusetts public schools, since MA has a very detailed formula for the calculation of funding for public schools. In order to measure student achievement, I first looked at a variety of factors, such as SAT scores, MCAS scores, and graduation rates. Since the SAT tends to be self selective and thus would introduce a missing data problem, I decided to use the MCAS results from 10th grade students, since the MCAS is mandated for all students, and thus does not sustain any missing data bias. This data is in the form of the percentage of students who have been deemed advanced or proficient in the results of the exam. Thus, this thus further restricts us to public high schools in Massachusetts. The data regarding school statistics and makeup was collected from a series of reports by the Massachusetts Department of Education. The dataset was then augmented with reports on the median income of every MA town, collected from the United States Census, with reports on the median property value of every town, collected from the Massachusetts Association of Realtors, the Cost of Living of every town in Massachusetts, and the political affiliation of every town, measured by the proportion of Democratic votes in the 2020 election. Since school sizes may vary greatly and thus total funding is not comparable, we will be using spending per pupil as our measure of school funding, since this is what is used by the state in order to compute the funding of schools. The median per pupil spending in Massachusetts is $14,116, and thus we define as our treatment variable whether a school has spend above $14,116 per pupil during the academic year.

3.2 Graph Elicitation

In order to determine the causal effects of school funding on student achievement, we first need to construct an ADMG of the causal relations between the covariates and potential causes of all variables. Since we have a very large list of potential variables, an automatic graph creation algorithm would be highly inefficient, and in the field of public policies it may also be very inaccurate. Thus, the determination of causality between variables which determines the presence of an edge in the ADMG as described in the preliminary section - will be done through a thorough review of the previous literature on the subject. We will restrict ourselves to papers which imply causal relations between the variables, as papers which only admit to correlation will not give us the necessary priors to assume the conditional independence implied by causation.

Since I identified 16 variables to be essential to include in the DAG, and since the possibility of both a directed and undirected edge in between each node allows for a total of 345 different edges, I have included literature reviews for only those edges I deemed most important to the causal estimation. The rest of the edges can thus be assumed to be included/excluded due to either absurdity or sensibleness. The incoming edges to MCAS Scores were found to be class size, student ethnicity, teacher characteristics (Tow, 2006), socio-economic status (parent income), and native language (Hampden-Thompson and Johnston, 2006). All of the edges incoming to School Funding were found through a review of the formula that the state of Massachusetts uses to distribute funding to schools. These factors are Special Education Needs, Native Language of students, student ethnicity, average town property value, average parent income, and average cost of living. Furthermore, the edge between Political Affiliation and School Funding was included because towns may choose to use more funding than the state has deemed necessary for them, choices which usually fall along party lines (Russo et al., 2015). Deng (2020) found that native language had an impact on the earnings, and it is also known that race/ethnicity has an impact on income in the US (Akee et al., 2019). For lack of collected metric on teacher quality in the available data, I used teacher salary as a proxy for teacher quality, represented by a bidirectional edge between teacher salary and teacher quality, since increases to teacher salary has been found to increase the quality of the teachers (Leigh, 2012).

There is no edge from MCAS Scores to property value since Clapp et al. (2008) found that test scores have no impact on the property value of the surrounding area. There is no edge from MCAS Scores to teacher quality, since it has been found that teachers do not choose their schools based on the achievement of students, and that quality teachers prefer to stay at their school than transfer to one that may be considered "better" (Boyd et al., 2011). Figure 1 shows the final graph achieved through this elicitation. Variables for School Resources, Teacher Quality, and Home Life have been included in the graph despite being unmeasured variables since they provide clarity to the relationships between the rest of the variables.

After graph elicitation, I tested every single conditional independences implied by the graph that we found using an odds ratio test, and found 95% confidence intervals using a 1000 sample bootstrap with replacement.

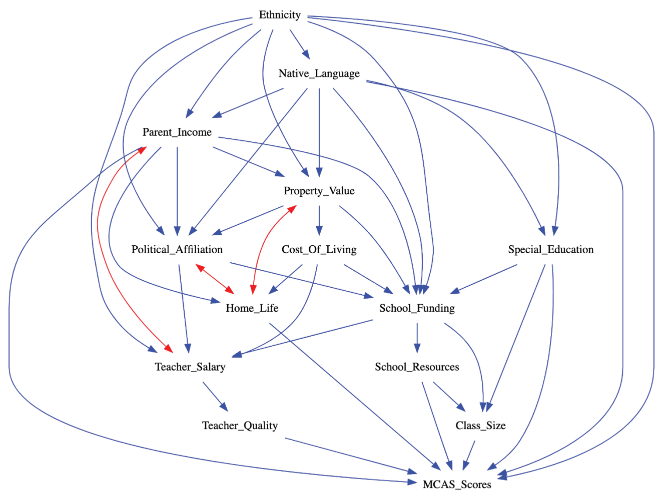

Figure 1: Original DAG derived from the graph elicitation

3.3 Casual Identification/Estimation

I will be computing the average causal effect (ACE) of school funding on the MCAS Scores of students. In order to do this, I will be using the backdoor criterion, originally conceived by Pearl (1995). Since the data does not come from randomized controlled trials, the backdoor criterion allows us to adjust for confounders in our attempt to isolate the ACE. The criteria denoted by the backdoor criterion to identify a set of conditioning variables Z are that the variables in Z cannot be descendants of A (School Funding), and that the variables in Z block all paths from Y (MCAS Scores) to A which end with an edge incoming to A, either a directed or bidirected edge. Using these criteria, we can identify the variables that we should include in Z, which would be the set {Special Education, Native Language, Ethnicity, Political Affiliation, Property Value, Parent Income, Cost Of Living}. This set is also an optimal minimal set, since it includes only factual parents of School Funding in the set, and since removing any one of these variables would cause the set to stop being a valid backdoor set. Thus, this set also provides us with the highest precision estimator relative to all of the other possible backdoor adjustment sets.

Using this backdoor set Z, I then computed the ACE using AIPW (Augmented Inverse Probability Weighting) method, which states that given a valid backdoor adjustment set Z on A and Y , the following equation holds true

$$ACE = E[Y(a) - Y(a')] = E\Big[\frac{\mathbb{I}(A=a)}{p(A|Z)} * (Y - \mathbb{E}[Y|A,Z]) + \mathbb{E}[Y|A=a,Z]\Big]$$

$$-E\Big[\frac{\mathbb{I}(A=a')}{p(A|Z)} * (Y - \mathbb{E}[Y|A,Z]) + \mathbb{E}[Y|A=a',Z]\Big]$$

We compute ACE using the Python Ananke package specified by Bhattacharya et al. (2020). The propensity score model used for $p(A|Z)$ would be a logistic regression, since we turn school funding into an indicator variable which is 1 when the average funding per pupil is above the mean funding in the state ($14,116), and 0 otherwise. Our model for E[Y |A,Z] is a regular regression model. This allows our estimates of the ACE using the AIPW to be double robust, which means that the ACE computed will converge to the truth when one of the following is true:

- The model p(A|Z) is correctly specified
- The regression model for E[Y |A,Z] is correctly specified.

In order to compute the confidence intervals for our ACE we ran 10,000 bootstrap samples with replacement and an alpha level of α = 0.05. The confidence intervals settled down after around 5,000 bootstrap samples, and thus 10,000 was deemed enough to produce significant intervals.

4. Results

In Table 1 (Page 16) we see all of the odds ratios associated with the DAG we elicited for the data. A majority of the confidence intervals have values very close to 1 with narrow confidence intervals, which leads us to believe that the DAG we created could be the true DAG for the data. Special Education   Teacher Salary| Ethnicity, Political Affiliation has a very small odds ratio ($2.7*10^{-17}$), but because of the extremely wide confidence interval range ($1.2*10^{-13}$,$8.5*10^{47}$), which could be explained by the small size of the dataset, since we are only looking at Massachusetts public high s⊥⊥ools, and the large amount of variables that we are conditioning on. However, the rest of the odds ratios are within the defined [0,1] range with minimal variation. Furthermore, the edges implied by removing the conditional independences which are computed to have large amounts of variation, such as the one described above, have found to have not effect on the backdoor adjustment set. Thus, despite the 3 variables with large ranges for the odds ration, we can continue with the DAG that we elicited in good faith. Thus, we are under the working assumption that the graph we are using is the true causal DAG, and that any edges which may have been added or omitted which should have been in the causal graph have no effect on the analysis done.

The ACE found exemplifies how many more students are deemed to have passed the MCAS proficiency level based when the school has above average funding, compared to the counterfactual. Our results show that for schools with above average funding, 2.17% more students are deemed proficient, with a confidence interval of (0.11,4.56). This means that we can say with significance that school funding Figure 2 shows a graph of student achievement on the MCAS exam plotted against the per pupil spending of the school, colored using the median income of the town that the school is in. Looking at this figure we see the difficulty of a straightforward statistical analysis of the relationship between our treatment and outcome, since the data is stratified by so many different variables. We see here that all of the lower achieving schools are also schools in low income areas, where school funding has been increased due to government aid and support. In Figure 3 we see the student achievement again plotted against per pupil spending, but this time divided by ethnicity. We again see a similar division, with schools with more diversity corresponding to the schools in low income areas which see lower academic achievement.

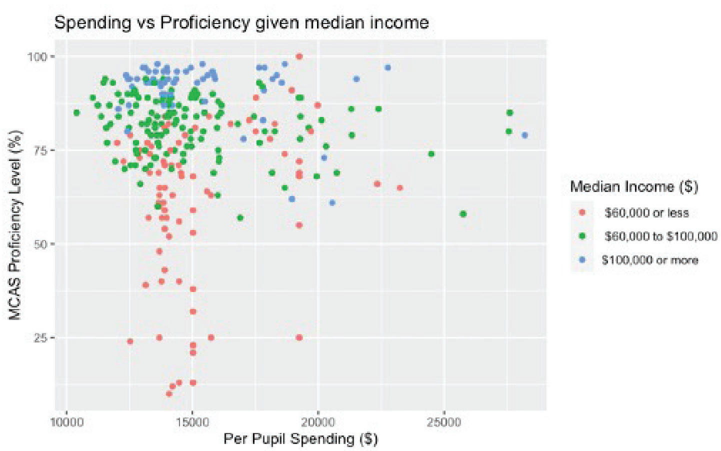

Figure 2: Proficiency Levels of students in a school given the per pupil spending, colored by median income level of the town

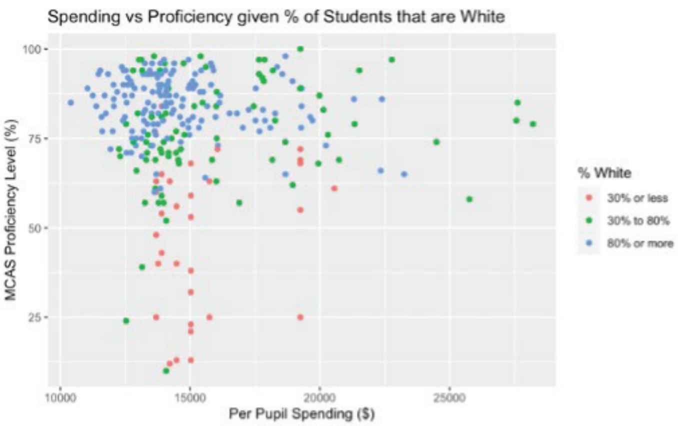

Figure 3: Proficiency Levels of students in a school given the per pupil spending, colored by the proportion of the school that is 'white'

Thus there is the importance of taking into account all of the other variables that may have a direct influence on student achievement, which is why taking a causal inference approach can help us make sense of the direct relationship. Although 2% is fairly small, with around 900,000 students in Massachusetts, this would account for about 18,000 students being below proficiency.

5. Discussion & Conclusion

The results of this paper are in line with the current range of critiques regarding the original Coleman Report published in the 1960's. Through the use of causal inference techniques I have found that decreases in the amount of per pupil funding provided by the school cause an increase in the amount of students that are deemed unprepared by the state, specifically when focusing on public high-school students in the state of Massachusetts. Thus, the public policy conversation regarding the equitable distribution of resources to schools in the U.S. should include the discussion of achieving equitable school funding, especially since better funded schools are often located in richer areas, since these areas have more funding that they can choose to allocate towards the schools, and thus the income inequality in the U.S. is directly related to the educational inequity.

In this paper I focused specifically on the state of Massachusetts, since MA has a very well defined school funding formula, and readily available data regarding its public schools. However, in the last few decades Massachusetts has taken steps towards achieving a more equitable distribution of funding in the state, which elicited the creation of the Chapter 70, which is what is used now to fund public schools (Berger and McLynch, 2006). Thus, Massachusetts is ahead of many of the other states in the U.S, which have done very little to combat the inequity in schools since the 1960s. Furthermore, Massachusetts is a fairly homogeneous state in regard to income and population size with respect to many of the other states in the U.S, where some rural schools may have less than 100 students. Thus, an analysis of the causal effect of funding on student achievement in states with less defined funding formulas and in states with a greater amount of income inequality would increase the evidence for federal reforms regarding school funding.

References

Randall Akee, Maggie R Jones, and Sonya R Porter. Race matters: Income shares, income inequality, and income mobility for all us races. *Demography*, 56(3):999–1021, 2019.

Noah Berger and Jeff McLynch. Public school funding in massachusetts. *Massachusetts Budget and Policy Center*, 2006.

Rohit Bhattacharya, Razieh Nabi, and Ilya Shpitser. Semiparametric inference for causal effects in graphical models with hidden variables. *arXiv preprint arXiv:2003.12659,* 2020.

Leah Boustan, Fernando Ferreira, Hernan Winkler, and Eric M Zolt. The effect of rising income inequality on taxation and public expenditures: Evidence from us municipalities and school districts, 1970–2000. *Review of Economics and Statistics*, 95(4):1291–1302, 2013.

Don Boyd, Hamp Lankford, Susanna Loeb, Matthew Ronfeldt, and Jim Wyckoff. The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools. *Journal of Policy Analysis and Management*, 30(1): 88–110, 2011.

John M Clapp, Anupam Nanda, and Stephen L Ross . Which school attributes matter? the influence of school district performance and demographic composition on property values. *Journal of urban Economics*, 63(2):451–466, 2008.

James S Coleman. Equality of educational opportunity. *Integrated education,* 6(5):19–28, 1968.

Ruyue Deng. The effect of the native language on immigrants. 2020.

Douglas B Downey and Dennis J Condron. Fifty years since the coleman report: Re-thinking the relationship between schools and inequality. *Sociology of Education*, 89(3): 207–220, 2016.

Gillian Hampden-Thompson and Jamie Johnston. Variation in the relationship between non school factors and student achievement on international assessments. 2006.

Dongsook Han, Sangheon Kim, Heeran Park, and Hosung Sohn. Promoting student achievement in high school using school funding: evidence from quantile regression discontinuity design. *Asia Pacific Education Review*, 22(2):193–208, 2021.

Daniel Kreisman and Matthew P Steinberg. The effect of increased funding on student achievement: Evidence from texas's small district adjustment. *Journal of Public Economics*, 176:118–141, 2019.

Julien Lafortune, Jesse Rothstein, and Diane Whitmore Schanzenbach. School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, 10(2):1–26, 2018.

Andrew Leigh. Teacher pay and teacher aptitude. *Economics of education review,* 31(3): 41–53, 2012.

Florence Neymotin. The relationship between school funding and student achievement in kansas public schools. *Journal of Education Finance,* pages 88–108, 2010.

Rachel R Ostrander. School funding: Inequality in district funding and the disparate impact on urban and migrant school children. *BYU Educ. & LJ*, page 271, 2015.

Judea Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 1995.

Charles J Russo, William E Thro, and Frank M Batz. Financing education: An overview of public school funding. *School Business Affairs*, 81(10), 2015.

Charlene Tow. The effects of school funding on student academic achievement. *Unpublished undergraduate thesis. University of California, Berkeley, Berkeley*, CA, 2006.

Conditional Independence Testing		
Conditional Independence	OR	Confidence Interval
Ethnicity ⊥⊥ Cost Of Living $\mid$ Property Value	1.03	(0.986, 1.078)
Ethnicity ⊥⊥ Class Size $\mid$ Political Affiliation, Funding	0.986	(0.966, 1.007)
Native Language ⊥⊥ Cost Of Living $\mid$ Property Value	0.98	(0.927, 1.033)
Native Language ⊥⊥ Class Size $\mid$ Funding, Political Affiliation	1.025	(0.999, 1.051)
Native Language ⊥⊥ Teacher Salary $\mid$ Funding, Political Affiliation, Cost Of Living, Ethnicity	$2.7 * 1-^{17}$	$(1.2 * 10^{-13}, 8.5 * 10^{47})$
Income ⊥⊥ Cost Of Living $\mid$ Property Value	1	(1.00006, 1.00035)
Income ⊥⊥ Special Education $\mid$ Native Language, Ethnicity	0.999	(.9999, 1.00001)
Income ⊥⊥ Class Size $\mid$ Political Affiliation, Funding	1	(0.999, 1)
Property Value ⊥⊥ Teacher Salary $\mid$ Funding, COL, Political Affiliation, Ethnicity,	0.996	(0.984, 1.005)
Property Value ⊥⊥ Special Education $\mid$ Ethnicity, Native Language	0.999	(0.999, 1)
Property Value ⊥⊥ Class Size $\mid$ Funding, Political Affiliation	1	(1, 1)
Special Education ⊥⊥ Cost Of Living $\mid$ Funding, Political Affiliation	0.993	(0.945, 1.033)
Special Education ⊥⊥ Teacher Salary $\mid$ Ethnicity, Political Affiliation	$1.3 * 10^{5}$	$(3.6 * 10^{-22}, 8.5 * 10^{27})$
Cost Of Living ⊥⊥ Political Affiliation $\mid$ Property Value	1.002	(0.998, 1.005)
Cost Of Living ⊥⊥ Class Size $\mid$ Funding, Political Affiliation	1.022	(1.008, 1.039)
Political Affiliation ⊥⊥ Special Education $\mid$ Ethnicity, Native Language	0.872	(0.026, 29.979)
Political Affiliation ⊥⊥ Class Size $\mid$ Funding, Political Affiliation	1.927	(0.696, 5.549)
Teacher Salary ⊥⊥ Class Size $\mid$ Funding, Political Affiliation	1.0001	(1.00005, 1.001)

Table 1: The conditional independences implied by the DAG and their respective odd ratio