OHIO EDUCATION RESEARCH CENTER

Making Research Work for Education

The Impact of Ohio STEM High Schools on Student Achievement

Stéphane Lavertu, PhD, The Ohio State University
Jennifer Gnagey, PhD, Weber State University

#FR-2015-13
STEM Education Initiatives
oerc.osu.edu

OHIO EDUCATION RESEARCH CENTER

The Ohio Education Research Center (OERC) is a **COLLABORATIVE** of Ohio-based researchers from six universities (Case Western Reserve University, Miami University, Ohio University, The Ohio State University, University of Cincinnati, and Wright State University) and four research institutions (Battelle, Battelle for Kids, Community Research Partners and Strategic Research Group) The founding partners coordinate the work of the OERC through the Governance Committee and three standing committees (Research Agenda, Data, and Outreach). Membership includes key participants from State of Ohio agencies and partner organizations. Administratively, the OERC reports to the State of Ohio through the Policy Council. The OERC is headquartered at The Ohio State University.

The **MISSION** of the OERC is to develop and implement a statewide, preschool-through-workforce research agenda addressing critical issues of education practice and policy. The OERC identifies and shares successful practices, responds to the needs of Ohio's educators and policymakers, and signals emerging trends. The OERC communicates its findings broadly, through multiple platforms and networks, producing materials, products and tools to improve educational practice, policy and outcomes.

The **VISION** of the OERC is to be the source for cutting edge knowledge and resources regarding education and training for Ohio's educators, policymakers and community leaders creating a dynamic cycle of research and practice where the needs of practitioners drive the research agenda and high-quality research has a rapid impact upon practice in the field.

Core **FUNDING** for the OERC is provided by the Ohio Department of Education. Additional funding comes from the Ohio Department of Job and Family Services in collaboration with the Ohio Board of Regents.

PREFACE

BY David L. Burns, Director of STEM Innovation Networks, Battelle Memorial Institute

On behalf of the Ohio STEM Learning Network, we wish to thank Stephane Lavertu and The Ohio State University for their work on the STEM (Science, Technology, Engineering and Mathematics) school study and the report they produced. We found Dr. Lavertu to be professional, courteous, and tenacious in his efforts to track students both inside and outside of STEM schools – no small task given the quality of data at his disposal and the multiple hoops necessary to jump through to work with secondary school student populations. By their nature, STEM schools receive a lot of attention from both local and national researchers. Organizations from Jobs for the Future to national foundations including the Bill & Melinda Gates have studied STEM schools. And they should. STEM schools are designed intentionally to be studied.

Ohio was one of the first states in the country to systemically design new STEM high schools based on competing in today's job market.

Each school is an experiment – specifically designed for research and development in STEM education in their local context. While each school was developed on a set of common design principles, the framework was flexible enough to allow for schools to be customized, meeting the needs of its local environment. Therefore, while the STEM school in Cincinnati was developed in the same manner as the STEM school in Cleveland, the differences between the two are more apparent than their similarities. By design, this holds true for all of Ohio's STEM schools. It was, and is, our belief that one model does not fit all. Schools should reflect their community. The more models we have, the better. This way, the lessons we learn from the variety of STEM schools allow us greater opportunity to improve the quality of education for all students in Ohio.

Which brings us back to Dr. Lavertu's report – (spoiler alert) that tells us that STEM schools do not prepare students as well as traditional schools do for the 10th grade Ohio Graduation Test. At first blush, the results are disconcerting. How could schools designed to prepare kids for college and career not do as well as traditional schools in preparing kids to meet the OGT? Then came the blinding flash of the obvious – STEM schools don't prepare kids to perform well on the OGT; traditional schools do. STEM schools prepare kids to get "through" the OGT as quickly as possible to get to the things that our schools, and more importantly, our communities, think matter – ACT/SAT scores, college coursework, and internships. STEM schools are focused on a different set of standards that have outcomes that we believe benefit students – not state report cards. As much as we would enjoy a vigorous debate on the merits of high stakes testing, standards and accountability, STEM schools were simply designed with a different end in mind than traditional high schools.

Metro Early College High School, Ohio's first STEM school, views its mission as being a "small school with a big footprint" and that element of its design was deliberately injected into the DNA of all of the STEM schools that followed. Metro is the most mature of the STEM schools in Ohio, founded in 2006, and its achievement results are impressive. Even more impressive considering that students are admitted using a first come, first served lottery process.

- 100% Graduation Rate
- 24.2 ACT composite score
- Average of 33 hours of Ohio State College credits earned per student

While all of Ohio's STEM schools have not yet matched Metro's results, they were designed with the same end goals as their objective. These are the kind of results they hope to achieve. Does this mean we ignore the report's results? No. As mentioned, at its core, the Ohio STEM Learning Network is a network of research and development schools. We welcome the findings and will look more closely at what we do and why we do it. Will we make a severe adjustment in our curriculum design and restructure our methods of teaching and learning hoping for higher levels of achievement on the OGT? No. We will look at what makes the most sense in the preparation of our young students to prepare them to be young adults. In this sense, we are very traditional – this is what it means to be a STEM school. We try, we learn, we try again.

ABSTRACT

Ohio government and private, non-profit organizations have invested significant resources to establish inclusive science, technology, engineering, and mathematics (STEM) high schools in Ohio. These schools feature non-selective admissions and, thus, have the potential to expand interest and aptitude in STEM fields among students who otherwise might not have access to a STEM-focused curriculum. This report presents the results of a study evaluating the impact on student achievement of six Ohio STEM high schools established since 2006. Using a variety of statistical modeling techniques—including multiple methods for matching STEM-school students to similar non-STEM-school students—the analysis compares the performance on 10th-grade mathematics, science, reading, and social studies achievement tests of students who attended these STEM high schools for two years to similar students who spent those two years in traditional district high schools.

The results suggest that, overall, attendance at these STEM high schools had a negative impact on student achievement across all four subjects—particularly the non-STEM subjects, reading and social studies. Further analysis indicates that the negative overall estimates in math and science are primarily attributable to students attending Hughes High School in Cincinnati and MC^2 in Cleveland. They also indicate that Dayton Regional STEM School's estimated positive impact on science achievement (a substantively significant impact corresponding to an extra 136 days of learning) is the only statistically significant positive effect across all schools and subject areas. Finally, the results do not indicate that schools open for a number of years are improving over time. We warn, however, that these results alone should not be used to determine the value of Ohio STEM high schools, particularly because standardized exams do not necessarily capture all outcomes of interest, such as higher-order problem-solving skills, educational attainment, and participation in the STEM workforce.

Funding for this report was provided by the Ohio Department of Education.

The Ohio Education Research Center would like to thank the following individuals who helped make this research possible:

Project Team
Stéphane Lavertu, Ph.D., The Ohio State University
Jennifer Gnagey, Ph.D., Weber State University

Questions regarding this report should be directed to Dr. Stéphane Lavertu at:
lavertu.1@osu.edu

CONTENTS

Executive Summary 2
I. Introduction 4
II. Background 6
III. Data 7
IV. Methods 10
V. Results 11
What is the impact of STEM school attendance by subject? 11
Does the impact vary across schools? 12
Are schools improving with each successive cohort? 13
VI. Conclusion 14
VII. References 15
Appendix A: Statistical Model 17
Appendix B: Matching Analyses 18
Appendix C: Instrumental Variable Analysis 20

EXECUTIVE SUMMARY

There is a widespread belief that investing in science, technology, engineering, and mathematics (STEM) education will improve students' educational achievement and attainment in STEM subjects, leading to better paying jobs and economic growth. Accordingly, since 2006 Ohio government and private nonprofit organizations have invested significant resources to establish inclusive STEM high schools, which feature non-selective admissions and, thus, have the potential to expand interest and aptitude in STEM fields among students who otherwise might not have access to a STEM-focused curriculum. But there is little systematic research that examines how these schools operate, and there is even less research that examines their impact on student educational and economic outcomes.

This report presents the results of a study evaluating the impact on student achievement of six Ohio STEM high schools established since 2006. These inclusive STEM high schools include two independent STEM schools—Dayton Regional STEM High School and Metro Early College High School—as well as four district schools—eSTEM Academy, Health Sciences & Human Services Academy ($(HS)^2$), MC^2 STEM High School, and Hughes STEM High School. The analysis compares the achievement on the 10th grade Ohio Graduation Test of students who attended the six STEM schools in grades 9 and 10 to similar students who attended a traditional public school during those two years. Importantly, the student-level analysis accounts for prior academic achievement and demographic characteristics to capture differences between students attending STEM schools and those attending traditional public schools. Additionally, to test the robustness of the results, we employed statistical matching techniques to create a variety of control groups for comparison and estimated models using a variety of specifications and samples.

The results indicate the following:

- Overall, attending an Ohio STEM high school had a negative effect on student achievement in math, science, reading, and social studies. These negative effects were modest in math and science, but they were substantively significant in reading and social studies.

- Analyses disaggregated by school indicate that the negative estimates in math and science are primarily attributable to Hughes High School in Cincinnati and MC^2 in Cleveland, and that Dayton Regional STEM School's positive achievement effect in science (a substantively significant positive impact corresponding to an extra 136 days of learning) is the only statistically significant positive effect across all schools and subject areas.

- Recognizing that schools can take a few years to establish themselves and realize their potential, the analysis also examines whether schools with multiple tested cohorts of students showed improvement over time. The results indicate that there generally were no positive trends in achievement in the STEM subjects, and two schools had negative trends in those subjects. The lack of positive time trends for schools with multiple tested cohorts suggests that those schools' null or negative results are not attributable to including initial cohorts in the analysis.

The purpose of Ohio STEM high schools is broadly to provide students with the opportunity to follow a STEM-focused curriculum, develop practices that promote student interest, achievement, and attainment in STEM fields, and, ultimately, strengthen the STEM workforce. Thus, achievement on standardized exams such as the Ohio Graduation Test—which, in particular, may not capture well achievement growth for students at the high-end of the achievement distribution—likely captures just a subset of what policymakers and other stakeholders hope to get from these STEM schools. Additionally, because some of these schools are in their early years of operation, it is important to defer judgment on achievement effects until those schools are completely established. However, this analysis suggests that, so far, Ohio STEM high schools as a group are not generating significant returns in terms of student knowledge and skills captured on standardized exams.

I. INTRODUCTION

There is a widespread belief that quality science, technology, engineering, and mathematics (STEM) education is the key to securing America's future. Political actors and policymakers at all levels of the U.S. government increasingly advocate policies meant to expand the scope and quality of the STEM workforce in order to meet the needs of the U.S. economy. In particular, many assert that policies promoting STEM education will improve students' educational achievement and attainment in STEM subjects, leading to better paying jobs and overall economic growth. Accordingly, since 2006 Ohio government and a number of private, non-profit organizations at the state and local level have helped launch STEM schools, many of which are "platform" schools that serve as a foundation for the Ohio STEM Learning Network (OSLN). These OSLN platform schools are meant to serve as laboratories for STEM education and disseminators of best practices. They are also inclusive STEM schools with non-selective admissions and, thus, have the potential to expand interest and aptitude in STEM fields among students who otherwise might not have access to a STEM-focused curriculum.

These platform schools hold significant promise for improving STEM education, promoting student interest and achievement in STEM fields, and, ultimately, expanding the STEM workforce. They also are widely considered to be a testament to what can be accomplished when public and private stakeholders join together to address public problems. But the establishment of inclusive STEM schools is a recent phenomenon. There is very little systematic research that examines how these schools operate, and there is even less research that examines their impact on student educational and economic outcomes. Indeed, there is only one study of which we are aware that examines the impact of STEM-school attendance on student-level outcomes using a rigorous statistical strategy, and this study focuses on STEM schools in New York City. An evaluation of Ohio's platform STEM schools is particularly important due to these schools' intended role as disseminators of best practices in STEM education.

This report presents the results of the first study to estimate the impact of six Ohio STEM platform schools on the academic achievement of the students who attended them. These inclusive STEM high schools include two independent STEM schools—Dayton Regional STEM High School and Metro Early College High School—as well as four district schools—eSTEM Academy, Health Sciences & Human Services Academy ($(HS)^2$), MC^2 STEM High School, and Hughes STEM High School. The analysis employs student-level data from the Ohio Department of Education and rigorous statistical techniques to compare the performance on 10th-grade mathematics, science, reading, and social studies achievement tests of students who attended the platform schools for two years (in grades 9 and 10) to similar students who spent those two years in traditional district high schools. Thus, it provides a first look at the academic benefits of these inclusive STEM high schools.

This study is not a summative evaluation of the academic value of these STEM schools. Most of the results presented below are based on these schools' first few years of operation, and research clearly indicates that such newly established schools typically take a few years to realize their full potential. Additionally, the analysis considers only a small subset of the outcomes of interest to policymakers—that is, the knowledge and skills captured by Ohio's 10th grade graduation test. It does not examine achievement in students' junior and senior years of high school, students' post-secondary achievement and attainment, or the ultimate outcome of interest: student employment in the STEM fields.

Nevertheless, this study is an important first step in evaluating the performance of STEM high schools in Ohio—particularly because research indicates that many of the educational and economic outcomes of interest are correlated with standardized test scores such as those examined in this study. Additionally, its results hold using rigorous statistical techniques meant to address bias associated with student self-selection into these schools. In the sections below, we provide some background on STEM education and, in particular, inclusive STEM high schools; review the study's data and methodology; present the results; and conclude by discussing the study's implications.

II. BACKGROUND

There have long been calls to promote U.S. STEM education in the name of economic prosperity and national security—most notably in response to the U.S.S.R.'s launch of sputnik in 1957 and since the publication of the National Commission on Excellence in Education's 1983 report, *A Nation at Risk*. Yet, such calls seem to have increased in recent years. Mostly notably, the National Academies' 2007 publication *Rising Above the Gathering Storm* asserts the importance of addressing the eroding achievement of U.S. students in mathematics and science relative to students in other countries, as well as addressing racial and gender achievement gaps in these fields (National Academies, 2007). The publication of such reports has been accompanied by calls for the improvement of secondary STEM education. In particular, the federal America COMPETES Act of 2007 called for providing "assistance to the states for the costs of establishing or expanding public, statewide specialty schools for mathematics and science" (America COMPETES Act, 2007), and in 2010 a report by the President's Council of Advisors in Science and Technology called for the creation of 1,000 more STEM-focused high schools within the next decade (PCAST, 2010).

Specialized STEM high schools have been around since the early 1900s, and there was a notable jump in their numbers in the latter half of the 20th century (Atkinson et. al. 2007). Today, 28 states have public residential STEM high schools and there are numerous local STEM magnet schools in operation (Thomas and Williams, 2010). These schools generally have been selective and have catered to students who have interest and excel in the STEM fields. Since the early 2000s, however, there has been a rapid increase in non-selective, "inclusive" high schools that have open enrollment policies. These emerging schools seek to build interest and talent in STEM disciplines among students who do not necessarily have exceptional interest or talent, thereby potentially expanding the STEM workforce pipeline and addressing gender and racial gaps in STEM interest and achievement (Means et. al. 2008). Research indicates that these inclusive STEM high schools serve more diverse student bodies than selective STEM schools, but they nevertheless have students who are more likely than non-STEM school students to report a desire to obtain a bachelor's degree and to pursue a career in science (Means et. al. 2008).

There is limited research on the impact of inclusive STEM schools on student achievement. Young et al (2011) evaluated the impact of the first 31 inclusive STEM schools established by the Texas High School Project's T-STEM initiative. They employed school-level data and used statistical techniques to match STEM high schools to observationally similar traditional public high schools. Their regression analysis indicates that T-STEM high schools were associated with greater achievement in math and science than the matched traditional public schools. To our knowledge, Wiswall et al.'s (2014) study of 30 selective and non-selective STEM high schools in New York City is the only study to use student-level data. They find positive impacts of STEM school attendance on some math and science exams. However, although their use of student-level data enables them to better account for selection bias, they do not account for students' prior achievement in science. As our analysis reveals, that is a potentially important omission.

Inclusive STEM schools across the U.S. are difficult to characterize because they generally do not operate according to a single organizational philosophy or model (Lynch et. al. 2013). The six schools evaluated in this report were established as inclusive STEM schools and have had ties to the Ohio STEM Learning Network (OSLN). Additionally, all have been designated a STEM school by the State of Ohio (see 33 Ohio Rev. Code 2007). But the schools are still quite different from one another. Two are independent schools, and four are district schools operated and funded by traditional public school districts. Some are newly established while others converted to STEM schools. More importantly, there is variation in their guiding philosophies and operations (e.g., see CEMSE, 2014) as well as in the students they serve.

III. DATA

The analysis focuses on STEM high schools that were part of the Ohio STEM Learning Network[1] and that had at least one cohort of students that had completed grades 9 and 10 by the end of the 2012-13 school year. As we indicate above, there is a lot of variation in how these schools operate. But all are state-recognized STEM schools that operated as OSLN platform schools and, thus, represent a more coherent group than the full set of Ohio schools that claim to have a STEM focus. Table 1 provides some basic information on these schools, including the number of student cohorts that had taken the 10th grade Ohio Graduation Test by the end of the 2012-13 school year—the latest year of our data.

Table 1. STEM High Schools in the Analysis

School Name	District	First year of operation	Student cohorts tested in Grade 10 (2012-13 or prior)	Grade span (as of 2012-13)
Metro Early College High School (Metro)	Independent	2006-07	6	9-12
Dayton Regional STEM School (DRSS)	Independent	2009-10	3	6-12
Hughes STEM High School (Hughes)	Cincinnati	2009-10	3	7-12
MC^2 STEM High School (MC^2)	Cleveland	2008-09	4	9-12
eSTEM Academy (eSTEM)	Reynoldsburg	2010-11	2	9-12
Health Sciences & Human Services Academy ($(HS)^2$)	Reynoldsburg	2011-12	1	9-12

The analysis primarily employs student-level administrative data that the Ohio Department of Education collected from the 2005-06 school year through the 2012-13 school year. In particular, the outcome of interest is student achievement on the 10th grade Ohio Graduation Test (OGT), and the treatment for which we estimate an impact is two years of attendance at a STEM high school. Thus, we are able to examine students who began attending these schools in 9th grade in 2011-12 or earlier. As Table 1 indicates, that enabled us to examine the impact of attending a STEM high school for between one and six student cohorts at each school.

---

[1] One STEM school affiliated with OSLN that we did not include is Linden-Mckinley STEM Academy in Columbus. We omitted this school due to recent controversies about their data and because the timing of the implementation of its STEM curriculum is unclear.

It is important to note that the grade spans have expanded since these schools were established. Most schools began with a single 9th grade cohort and then added grade levels over time.[2] It also is important to note that the OGT may not be an ideal measure to capture student achievement growth since grade 8, as the test contains a significant amount of content from prior grades and, in particular, may not enable one to capture achievement gains for students at the high end of the achievement distribution.[3] Results should be interpreted with these potential limitations in mind.

The student-level data we obtained from the Ohio Department of Education include student test scores in grades 8 and 10, demographic information (race/ethnicity, gender, and economic, English proficiency, disability, and gifted statuses), and student school and district attendance information.[4] Table 2 presents descriptive statistics for the restricted sample of students used to estimate the models presented in this report. Specifically, the sample is restricted to students for whom we had 10th grade test scores in all four subjects and who attended the STEM school or a STEM school's feeder districts for two consecutive years in grades 9 and 10.[5] The test scores are scale scores standardized by year, grade, and subject and reported in standard deviation units. Thus, a student receiving a scale score equal to the statewide average for that subject, year, and grade gets a value of zero, one who scores below average gets a negative value, and one who scores above average gets a value greater than zero. Finally, the table also reports the percent of the STEM school or feeder district students who fit into various demographic categories and the percent who previously attended charter schools.

Table 2 illustrates the significant differences between the STEM school students and their counterparts in traditional public schools, as well as the significant differences between the STEM students in each of the six schools. For example, the two independent STEM schools draw from numerous districts (that is why there are many feeder district students for comparison), and their students—as well as students who attended district schools MC^2 and eSTEM—have significantly higher prior test scores than their counterparts attending traditional public schools in their districts of residence. On the other hand, the students at Hughes—a district STEM school—tend to have lower 8th grade test scores, are more economically disadvantaged, are more likely to be African American, are more likely to be female, and

[2] This detail is particularly relevant for Dayton Regional STEM School (DRSS), which began with a 9th grade cohort in 2009-10 and then added new 8th and 9th grade cohorts the following year, for a full grade span of 8-10 in 2010-11. Thus, some 10th grade students in 2012-13 had been at the school for three years when they took the OGT. Nevertheless, we coded these students as having attended DRSS "high school" for two years.
[3] In particular, if OGT captures basic content that more advanced students covered in prior grades, then it will do a poor job of differentiating between advanced students. As Table 1 indicates, students attending Metro, DRSS, and eSTEM were relatively high achieving, and the analysis compares their achievement to similarly high-achieving students. If these students' coursework is focused on material that is more advanced than the basic OGT material, then analyzing test scores on the OGT may not allow us to detect differences between high-achieving STEM school and non-STEM school students. As the results indicate, however, this analysis was able to identify significant positive achievement effects among the relatively high-achieving students at DRSS. Additionally, this should be less of an issue for the other schools, whose students were average or on the low end of the achievement distribution.
[4] For one of the six STEM high schools, we also obtained admissions data on all students who applied to the school. These admissions data were necessary to identify all of the students who attended this particular school in its early years, as the school was technically a "program" at the time and state programmatic records were incomplete. Though it was technically a "program," the school was nonetheless a standalone school that had its own building.
[5] The descriptive statistics are similar for the full sample. A more complete table for the unrestricted sample is available upon request. Also note that the results presented below are based on an analysis that is further restricted to only those students for whom we have values of demographic variables and 8th grade test scores in math, science, and reading. The models on average capture three-fifths of all students attending the schools included in the analysis. Missing data was particularly an issue for Hughes and MC^2, as only 50 percent of their students enter the analysis. But missing data are only problematic if there is reason to believe that the difference in impact of STEM and non-STEM schools would be different if the analysis focused on those students with missing observations. It is also important to note that our statistical modeling strategy accounts for differences between students in the two sectors.

are less likely to be labeled as gifted than their feeder district counterparts. Finally, students attending independent STEM schools (Metro and DRSS), as well as district school MC^2, were more likely to have attended charter schools than their traditional district counterparts.

Table 2. Descriptive Statistics for Student Sample used in the Analysis

	Metro		**DRSS**		**Hughes**	
	STEM	Feeder	STEM	Feeder	STEM	Feeder
Mean standardized 8th grade math score	0.71	0.08	0.52	0.15	-0.55	-0.11
Mean standardized 8th grade reading score	0.71	0.06	0.54	0.14	-0.46	-0.03
Mean standardized 8th grade science score	0.65	0.03	0.71	0.07	-0.75	-0.25
Pct Female	0.46	0.50	0.38	0.51	0.62	0.55
Pct African-American	0.27	0.29	0.24	0.19	0.92	0.64
Pct Hispanic	0.04	0.04	0.00	0.02	0.01	0.01
Pct Other non-white	0.12	0.07	0.12	0.07	0.02	0.06
Pct Economic disadvantage	0.40	0.39	0.29	0.33	0.86	0.56
Pct Limited English proficiency	0.05	0.05	0.00	0.01	0.04	0.02
Pct Disability	0.06	0.11	0.07	0.11	0.20	0.16
Pct Gifted	0.68	0.34	0.22	0.23	0.08	0.25
Pct attended charter in 8th grade	0.05	0.01	0.32	0.02	0.05	0.04
Student Count	241	29,626	100	9,435	211	3,132
	MC^2		**eSTEM**		**$(HS)^2$**	
	STEM	Feeder	STEM	Feeder	STEM	Feeder
Mean standardized 8th grade math score	-0.29	-0.60	0.67	0.05	0.06	0.01
Mean standardized 8th grade reading score	-0.27	-0.52	0.57	0.09	0.13	0.07
Mean standardized 8th grade science score	-0.37	-0.77	0.86	0.20	0.30	0.20
Pct Female	0.45	0.54	0.36	0.53	0.78	0.57
Pct African-American	0.73	0.70	0.24	0.39	0.53	0.37
Pct Hispanic	0.12	0.12	0.02	0.01	0.02	0.01
Pct Other non-white	0.03	0.03	0.11	0.11	0.10	0.14
Pct Economic disadvantage	0.94	0.95	0.26	0.39	0.47	0.40
Pct Limited English proficiency	0.04	0.05	0.03	0.04	0.08	0.04
Pct Disability	0.11	0.18	0.04	0.12	0.04	0.11
Pct Gifted	0.27	0.15	0.43	0.19	0.27	0.23
Pct attended charter in 8th grade	0.12	0.03	0.00	0.01	0.00	0.01
Student Count	103	6,048	167	283	51	90

Note: The table presents descriptive statistics for the restricted sample of students used to estimate the models presented in this report. Specifically, the sample is restricted to students for whom we had 10th grade test scores in all four subjects and who attended the STEM school or a STEM school's feeder districts for two consecutive years in grades 9 and 10. Test scores are scale scores standardized by year, grade, and subject and are reported in standard deviation units.

IV. METHODS

The analysis compares the achievement of students who attended these STEM schools in grades 9 and 10 to the achievement of students who attended the traditional district schools that the STEM school students would have attended had they chosen to attend a traditional public school in their district of residence. We employed a number of statistical strategies to estimate differences in student achievement. However, all statistical strategies entailed the estimation of student-level growth in academic achievement. Specifically, all models include variables capturing students' 8th grade test scores in math, science, and reading. Accounting for prior test scores in this way accounts for differences in student ability and educational histories through grade 8 and, thus, to a significant extent accounts for differences between students that attended STEM and non-STEM schools. Indeed, accounting for prior test scores has been shown to eliminate selection bias in the estimation of teacher impacts on student test outcomes (Chetty et al, 2014; Dieterle, 2015; but see Rothstein, 2014). Additionally, the statistical models account for student gender, race, economic disadvantage, limited English proficiency, disability status, gifted status, test year, district, and whether or not a student attended a charter school in grade 8. (See Appendix A.)

The multitude of models we estimated differ primarily in terms of the students they compare. The results we report below are based on models that compare the achievement of each STEM high school's 10th grade students to 10th grade students in the districts from which that STEM school draws. For example, the independent STEM schools draw students from multiple urban and suburban districts, and the analysis includes all 10th grade students who attended each of those districts in grades 9 and 10. Additionally, the results reported below are limited to students for whom we had test scores in all four tested subjects on the Ohio Graduation Test. We restricted the sample in this way to ensure that we are making comparisons between similar samples of students across each tested subject. Importantly, the results are similar to those we obtained using models that were not restricted to students with test data in all subjects, as well as to the results of analyses for which we created comparison samples using multiple student matching techniques. (Appendix B describes the matching methods we used and Table B1 presents the results of the matching analyses based on the restricted sample of students.)

Finally, we used one additional statistical strategy to estimate the impact of the school that provided us with admissions data. A powerful strategy for estimating the causal impact of schools of choice is to compare the achievement of those who applied to attend but, by random chance, were not granted admission, to those who applied to attend and, by random chance, were granted admission (e.g., see Betts & Tang, 2011). We employed the lottery admissions data to estimate the causal impact of attending that STEM school using an instrumental variables approach. The results for that school were similar to those we report below, as well as to those obtained using the various matching techniques we describe in the appendix.

We were unable to obtain such admissions data from the other five schools. However, that the results for this particular school are similar lends some reassurance that the results we report identify the causal effects of the STEM schools. Additionally, academic research indicates that matching techniques such as those we use yield estimates similar to those one would get using admissions lotteries as an instrument (Bifulco, 2012), and matching methods are used in reputable studies examining the impact of schools of choice (e.g., see CREDO, 2014).

V. RESULTS

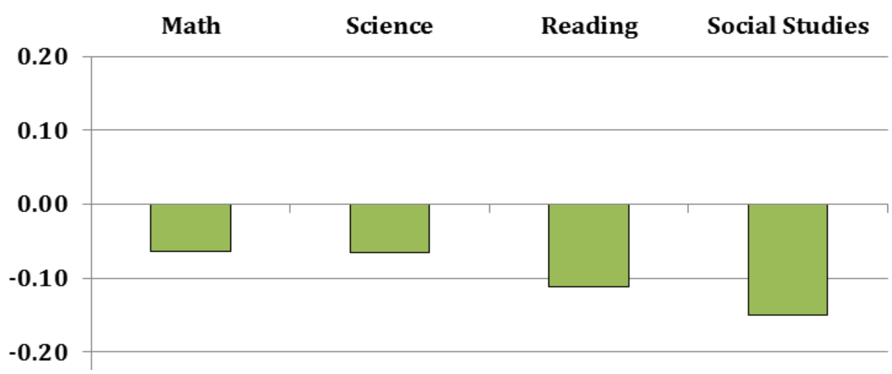
The results we present in this section capture the difference in achievement on 10th grade tests between students who attended the six STEM schools for two years to students who attended traditional public schools for those two years. The estimates are from models that account for students' 8th grade test scores in math, reading, and science; gender and race (African American, Caucasian, Hispanic, or other); disability, gifted, English proficiency and economic statuses; the district in which students reside; and whether or not students attended a charter school in 8th grade. In the interest of space, we report only the impact of STEM school attendance.

We report the results in standard deviation units. For example, if the estimated effect is -0.1, that indicates that students attending a STEM school scored 0.1 standard deviations worse on the test than similar students who attended traditional public schools. To give the non-technical reader a sense for the magnitude of these impacts, scholars sometimes convert their estimates into "days of learning," assuming that in a given school year students typically experience 180 days of learning. Using the estimates of yearly student achievement growth for grades 8-10 in math and reading (see Hill et al, 2008), one can characterize achievement effects of 0.05, 0.10, 0.15, and 0.20 standard deviations as roughly 40, 80, 120, and 160 days of learning, respectively. Thus, an achievement effect of 0.10 corresponds to roughly an extra 80 days of learning for STEM school students, and an effect of -0.10 corresponds to roughly 80 fewer days of learning for STEM school students.[6] Finally, we indicate whether results in the tables are statistically significant (at the conventional $p<0.05$ level for a two-tailed test) with solid bars. (The precise coefficient estimates, standard errors, and significance levels are presented in Appendix A, Table A1.)

What is the impact of STEM school attendance by subject?

Figure 1 presents results aggregated across all six STEM platform schools. The figure indicates that attending these STEM schools is associated with lower achievement in math (-0.06), science (-0.07), reading (-0.11), and social studies (-0.15); that these results are all statistically significant; and that the negative effects are more pronounced for non-STEM subjects.

Figure 1. Impact of STEM high school attendance in standard deviation units



Note: The bars indicate the difference in achievement between students attending STEM schools and those attending traditional public schools in feeder school districts. A negative value indicates that students attending STEM schools perform worse on average than those not attending STEM schools. Solid bars indicate statistically significant effects at the conventional $p<0.05$ level for a two-tailed test.

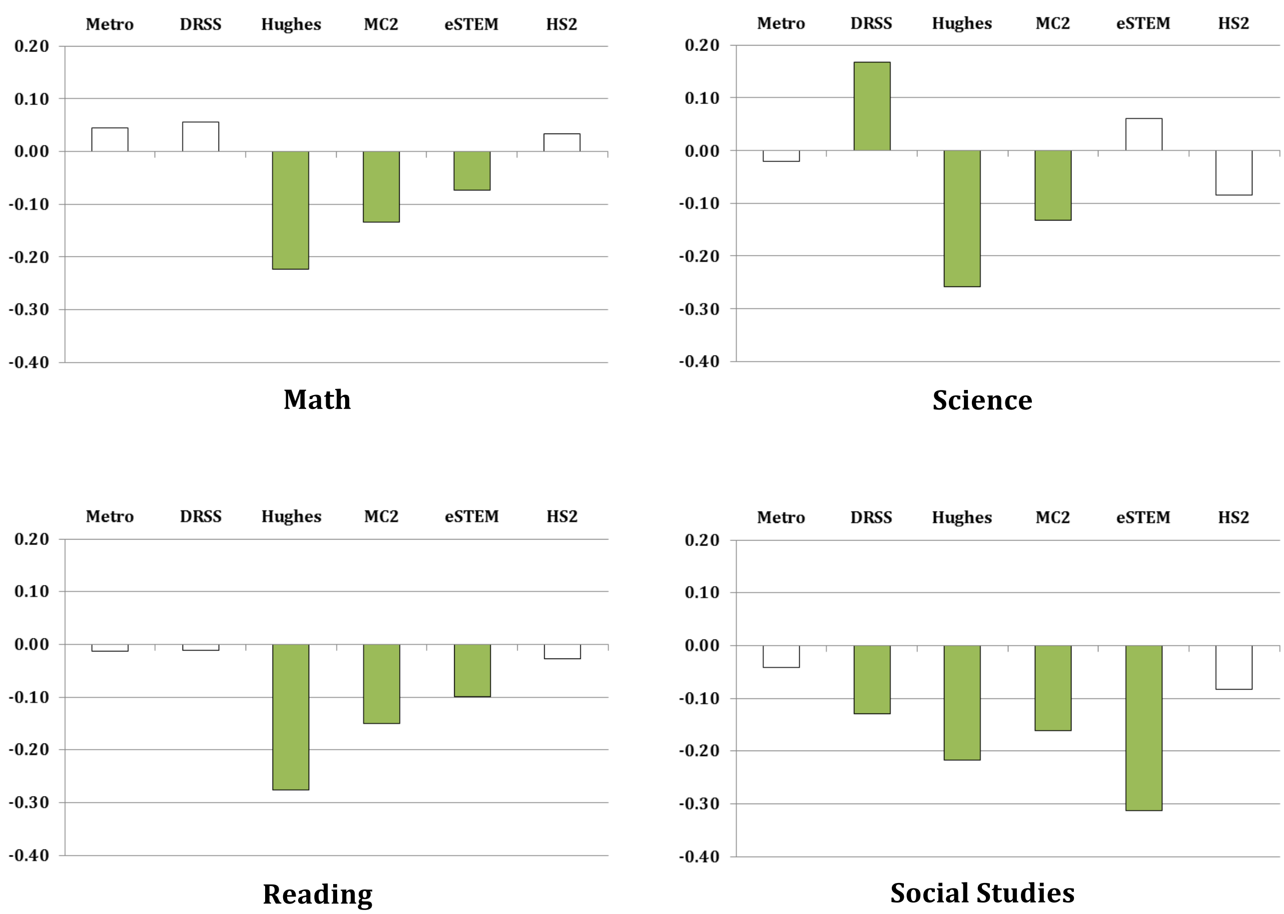
---

[6] Note that these effect sizes are larger than those presented in the CREDO reports (e.g., CREDO, 2014) because average achievement gains are much smaller in later grades than they are in the earlier grades that are the focus of the CREDO reports.

Does the impact vary across schools?

Though they are all state-designated STEM schools and OSLN platform schools, treating all six high schools as if they represent a coherent educational intervention is highly problematic for a number of reasons. Some were in their early years of operation (particularly Reynoldsburg eSTEM and $(HS)^2$); some are district schools and others are independent; and, as we discuss above, all are meant to serve as laboratories and entail significant differences in their operations. Figure 2 disaggregates the estimated achievement effects by school.

Figure 2. Impact of STEM high school attendance in standard deviation units (by school)



Note: The bars indicate the difference in achievement between students attending STEM schools and those attending traditional public schools in feeder school districts. A negative value indicates that students attending STEM schools perform worse on average than those not attending STEM schools. Solid bars indicate statistically significant effects at the conventional p<0.05 level for a two-tailed test.

The results indicate that, across all tests and schools, only Dayton Regional STEM School is associated with a positive and statistically significant effect on science achievement. The effect of 0.17 standard deviations is also substantively significant, corresponding to roughly 136 days of learning. These results are tempered somewhat in the matching analyses (see Appendix B, Table B1). Besides DRSS's impact in science, the data do not allow us to say whether students attending independent STEM high schools do any better or worse than similar students who attended traditional public schools. On the other hand, the results indicate that attending Hughes in Cincinnati and MC^2 in Cleveland had significant negative effects across all four subjects. Reynoldsburg's eSTEM also has significant negative impacts in non-STEM subjects—but the negative effects in math and reading reported in Figure 2 are not entirely robust to the use of matched student samples (see Appendix B, Table B1).

Are schools improving with each successive cohort?

Judging the quality of these schools based on their first few years of operation is problematic, as it can take a few years for a school to establish itself. To address this concern, we examined trends in these schools' performance over time. There is no evidence of a positive and linear trend in science and math for eSTEM, Metro, and DRSS (though the results are somewhat volatile from year to year for the latter two). There is, however, evidence of a negative trend at Cleveland's MC^2. For example, whereas the first two MC^2 cohorts performed no better or worse than those in traditional public schools in math, the third and fourth cohorts performed far worse, with substantively significant effect sizes of -0.24 and -0.20. Similar results obtain for MC^2's performance in science.[7] Finally, the results indicate that Hughes High School in Cincinnati made significant improvements over time in math, but they also indicate a significant decline in science. For example, math achievement effects went from -0.31, to -0.23, and then to -0.14 for the last cohort examined. On the other hand, the achievement effects in science declined from -0.19, to -0.23, to -0.34—substantively and statistically significant negative effects.

[7] It is worth noting, however, that there were a lot of missing data from MC^2 in those first two years.

VI. CONCLUSION

The analysis of the achievement effects of Ohio STEM high schools compares the achievement on the 10th grade Ohio Graduation Test of students who attended the six STEM platform high schools in grades 9 and 10 to similar students who attended a traditional public school during those two years. The results indicate the following:

- Overall estimated effects across the six schools and student cohorts tested through the 2012-13 school year are negative across all subjects.
- These negative estimated impacts are substantively most significant for the two non-STEM subjects: reading and social studies.
- The negative overall results in math and science are primarily attributable to Hughes High School in Cincinnati and MC^2 in Cleveland.
- Dayton Regional STEM School's positive achievement effect in science—a substantively significant positive effect that corresponds to roughly 136 days of learning—is the only statistically significant positive effect across all schools and subject areas.
- The lack of positive time trends for schools with multiple tested cohorts suggests that those schools' results are not attributable to the analysis including initial cohorts, when these schools were just getting established.

We wish to reiterate, however, that OSLN STEM schools pursue a multitude of goals. Achievement on standardized exams such as the Ohio Graduation Test likely captures just a subset of the outcomes of interest to policymakers and other stakeholders. With respect to the limitations of the OGT in particular, future research should replicate this study with the new exams that Ohio is implementing, as these should enable one to capture differences between students who are at the low or high end of the achievement distribution. Nevertheless, this analysis suggests that, so far, Ohio STEM high schools as a group are not generating significant returns in terms of student knowledge and skills captured on the Ohio Graduation Test.

VII. REFERENCES

33 Ohio Rev. Code. (2007). Section 3326 et. seq. (2014).

America COMPETES Act. (2007). 42 U.S.C. Section 7381h (2007).

Atkinson, R. D., Hugo, J., Lundgren, D., Shapiro, M. J., & Thomas, J. (2007). *Addressing the STEM challenge by expanding specialty math and science high schools.* Washington, D.C.: The Information Technology and Innovation Foundation. Retrieved May 1, 2014, from http://www.itif.org/publications/addressing-stem-challenge-expanding-specialty-math-and-science-high-schools

Bifulco, R. 2012. Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison. *Journal of Policy Analysis and Management,* 31, 729-751.

Betts, J. R., & Tang, E. (2011). *The effects of charter schools on student achievement: a meta-analysis of the literature.* Seattle, WA: Center on Reinventing Public Education at the University of Washington Bothell. Retrieved May 1, 2014, from http://www.crpe.org/publications/effect-charter-schools-student-achievement-meta-analysis-literature

CEMSE (Center for Elementary Math and Science Education). (2014). *The CEMSE STEM school study.* Online Project Description. Chicago, IL: CEMSE at the University of Chicago. Retrieved May 1, 2014, from http://outlier.uchicago.edu/s3/

Chetty, R., Friedman, J., & Rockoff, J. (2014). Measuring the Impact of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review,* 104, 2593-2632.

Center for Research on Education Outcomes (CREDO). (2014). *Charter School Performance in Ohio* Stanford, CA: Stanford University. Retreived January 4, 2015, from http://credo.stanford.edu/pdfs/OH_Report_2014F.pdf

Dieterle, S., 2015, Cuarino, C.M., Reckase, M.D., & Wooldridge, J.M. (2015). How do principals assign student to teachers? Finding evidence in administrative data and the implications for value added. *Journal of Policy Analysis and Management,* 34(1), 32-58.

Hill, C.J., Bloom, H.S., Rebeck Black, A., & Lipsey, M.W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, 2(3): 172-177.

Lynch, S. J., Behrend, T., Burton, E. P., & Means, B. (2013). *Inclusive STEM-focused high schools: STEM education policy and opportunity structures.* Paper presented at the NARST 86th Annual Converence. Rio Grande, Puerto Rico.: National Association for Research in Science Teaching. Retrieved May 1, 2014, from http://ospri.research.gwu.edu/conference-presentations

Means, B., Confrey, J., House, A., & Bhanot, R. (2008). *STEM high schools: specialized science, technology, engineering, and mathematics high schools in the U.S.* Report prepared for the Bill and Melinda Gates Foundation for Project P17858. Menlo Park, CA: SRI International. Retrieved May 1, 2014, from http://www.sri.com/work/publications/stem-high-schools

National Academies. (2007). *Rising above the gathering storm: energizing and employing America for a brighter economic future.* Washington, D.C.: National Academies Press. Retrieved May 1, 2014, from http://www.nap.edu/catalog.php?record_id=11463

Rothstein, Jesse. (2014). Revisiting the Impacts of Teachers. UC-Berkeley Working Paper.

Subotnik, R. F., Tai, R. H., Almarode, J., & Crowe, E. (2013). What are the value-added contributions of selective secondary schools of mathematics, science, and technology? - preliminary analyses from a U.S. national research study. *Talent Development and Excellence, 5*(1), 87-97.

Thomas, J., & Williams, C. (2010). The history of specialized STEM schools and the formation and role of NCSSSMST. *Roeper Review, 32*, 17-24.

Wiswall, M., Stiefel, L., Schwartz, A.E., & Boccardo, J. (2014). Does attending a STEM high school improve student performance? Evidence from New York City. *Economics of Education Review,*40,93-105.

APPENDIX A: STATISTICAL MODEL

We estimated achievement growth models common in the education evaluation literature. These models feature lagged test scores that control for the pre-treatment effects of observable inputs and the time-constant effect of unobservable initial endowments that affect student achievement levels. Treatment effects are identified under the assumption of common trends conditional on observables. Specifically, we estimated the following model:

$$y_{ig} = \beta_0 + \lambda \boldsymbol{y}_{ig-2} + \beta \boldsymbol{x}_{ig} + \gamma^s \boldsymbol{D}_{ig}^s + \varepsilon_{ig}$$

where y_{ig} is student i's test score in 10th grade; $\boldsymbol{y}_{ig-2}$ is a vector of one or more 8th grade test scores; $\mathbf{D}_{\mathrm{ig}}^{\mathrm{s}}$ is a vector of dummy variables indicating attendance at each of the six schools and γ^{s} is a vector of school specific treatment effects; the vector $\boldsymbol{x}_{ig}$ represents observable student-level demographic characteristics and includes indicators for gender, race, eligibility for free or reduced price lunch, limited English proficiency, disability status, gifted status, charter school attendance in 8th grade, and cohort-by-district-of-residence fixed effects; and ε_{ig} is a random error term.

Table A1 presents the estimated STEM school effects from OLS regressions based on the broader sample of students. Robust standard errors are reported below the regression coefficients. Significance levels are based on two-tailed test: *p<0.05 or ^p<0.10.

Table A1. STEM HS estimates using "full" sample of students from feeder districts

	Math	**Science**	**Reading**	**Social Studies**
All STEM Schools	**-0.064***	**-0.066***	**-0.112***	**-0.150***
	(0.020)	(0.019)	(0.021)	(0.021)
Independent STEM School	0.047	0.033	-0.013	**-0.067***
	(0.031)	(0.028)	(0.028)	(0.032)
District STEM School	**-0.151***	**-0.145***	**-0.189***	**-0.216***
	(0.024)	(0.025)	(0.029)	(0.027)
Metro	0.044	-0.021	-0.013	-0.042
	(0.038)	(0.033)	(0.034)	(0.038)
DRSS	0.055	**0.167***	-0.012	**-0.129***
	(0.051)	(0.050)	(0.045)	(0.055)
Hughes	**-0.223***	**-0.259***	**-0.276***	**-0.216***
	(0.041)	(0.037)	(0.047)	(0.045)
MC2	**-0.135***	**-0.132***	**-0.150***	**-0.161***
	(0.041)	(0.052)	(0.061)	(0.046)
eSTEM	**-0.074***	0.061	**-0.099***	**-0.312***
	(0.041)	(0.042)	(0.044)	(0.042)
HS2	0.034	-0.085	-0.027	-0.083
	(0.072)	(0.062)	(0.069)	(0.073)
Treated students	873	873	873	873
Control students	48,524	48,524	48,524	48,524
Total students	49,397	49,397	49,397	49,397

APPENDIX B: MATCHING ANALYSES

We estimated all models using a variety of samples of students for comparison. The results presented above are based on models estimated using all 10th grade students in the six STEM schools and all 10th grade students in the traditional public schools in the STEM school students' districts of residence (provided that we have test scores across all four subjects for each student). This approach could be problematic if STEM school students are not representative of students in their districts of residence. In other words, the control group used in the analysis (i.e., all 10th grade students in STEM school students' districts of residence) may not be sufficiently similar to the treatment group (i.e., the STEM school students). There are a number of statistical matching techniques available to create control groups of non-STEM students who are similar to STEM school students according to students' observable characteristics—techniques to improve sample balance.

We employed a number of matching techniques to create treatment and control groups that are similar in terms of the student characteristics included as covariates in the model described above. Ultimately, we decided to employ comparison samples we created using propensity score matching methods. Propensity score matching consists of matching students based on their probability of attending a STEM school based on observable characteristic. Specifically, the approach entails estimating this probability (i.e., a propensity score) for each student within a cohort and district, matching students with similar propensity scores, and estimating the models based on the matched samples.

We created a number of student samples and tested each for balance in the characteristics between the groups of STEM and non-STEM students. Ultimately, to maximize sample balance, we focused on samples that we created using estimated probabilities we generated using probit models that feature quadratic specifications for 8th grade test scores, interactions between these test score variables, interactions between indicators of student demographic characteristics, and cohort-by-district-of-residence fixed effects. We estimated separate probit models for each STEM school.

After estimating the propensity scores, we matched students exactly using cohort and district of residence. Then, we created two matched samples for analysis using two different matching rules. We created one sample using 1-to-1 nearest-neighbor matching within calipers and with replacement. This entailed matching STEM students to non-STEM students in their cohort and district that had the nearest propensity score with a caliper equal to 0.25 standard deviations of the pooled propensity scores. We created the second sample using on 1:N radius (or caliper) matching with replacement. In other words, in this case STEM school students are matched to all non-STEM school students within the caliper.

After testing for balance, we estimated the regression model we describe above using weights in order to account for the imbalance in counts between treatment and control groups due to sampling with replacement.

Table A2 presents the coefficients estimated using the matched samples, as opposed to the "full" sample of students attending traditional public schools in the STEM schools' feeder districts. The table presents the estimated STEM school effects from OLS regressions based on the 1:1 "nearest neighbor" matching methods (odd-numbered columns) and radius matching methods (even-numbered columns). Each column presents the estimated effects from three separate regressions: 1) all STEM schools, 2) disaggregated by STEM school type, and 3) disaggregated by school.

Standard errors are report below the regression coefficients. Bolded coefficients are significant at the following levels for a two-tailed test: *p<0.05 or ^p<0.10

Table B1. STEM high school treatment effects estimated using propensity score matched samples								
	Math		**Science**		**Reading**		**Social Studies**	
	1:1 NN	**Radius**	**1:1 NN**	**Radius**	**1:1 NN**	**Radius**	**1:1 NN**	**Radius**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
All STEM Schools	**-0.09***	**-0.08***	**-0.09***	**-0.09***	**-0.10***	**-0.11***	**-0.20***	**-0.18***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Independent STEM School	0.01	0.02	-0.01	-0.01	0.00	-0.03	**-0.13***	**-0.10***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
District STEM School	**-0.15***	**-0.14***	**-0.14***	**-0.14***	**-0.16***	**-0.15***	**-0.25***	**-0.24***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Metro	0.02	0.01	-0.06	-0.07	0.01	-0.05	**-0.10^**	-0.07
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
DRSS	-0.04	0.04	0.12	**0.15^**	-0.02	0.02	**-0.23***	**-0.17^**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
Hughes	**-0.24***	**-0.21***	**-0.26***	**-0.25***	**-0.28***	**-0.27***	**-0.24***	**-0.21***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
MC2	**-0.12^**	**-0.17***	**-0.12^**	**-0.18***	**-0.14^**	**-0.19***	**-0.18***	**-0.21***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)
eSTEM	**-0.11***	-0.07	0.02	0.04	-0.06	-0.02	**-0.33***	**-0.33***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
HS2	0.03	0.03	**-0.17***	**-0.15^**	-0.02	0.01	-0.12	-0.08
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)
Treated students (weighted)	839	839	839	839	839	839	839	839
Control students (weighted)	839	839	839	839	839	839	839	839
Total students (weighted)	1678	1678	1678	1678	1678	1678	1678	1678

APPENDIX C: INSTRUMENTAL VARIABLE ANALYSIS

The lottery admissions data from one of the STEM high schools enabled us to compare the achievement of STEM and non-STEM students that had applied to attend the school. Specifically, we used lottery outcomes as an exogenous instrument for STEM school attendance. We then estimated the student growth model using Two Stage Least Squares (2SLS). The results were similar to those produced using the techniques described above.