Preliminary Effects of the School Improvement Grant Program on Student Achievement in Texas

David Dickey-Griffith

David Dickey-Griffith graduated from the McCourt School of Public Policy in May 2013. Previosly, David worked for Congressman Earl Blumenauer of Oregon's 3rd Congressional District, as well as the 2008 campaign of US Senator Jeff Merkley. David is currently a high school social studies teacher at Perry Street Preparatory Public Charter School in NE Washington, DC.

Abstract

This paper uses school-level data from the state of Texas to test whether receiving a School Improvement Grant (SIG) has led to higher graduation, completion, or dropout rates and/or increased student achievement, as measured by standardized tests in reading, math, science, social studies, and writing. Part of this analysis tests to see if the effects of the program vary for urban versus rural schools, charter versus non-charter schools, or between demographic subgroups. This paper's results suggest that in its first year of implementation, SIG had a negative effect on student achievement at elementary and middle schools across almost all subjects and subgroups and had little effect on achievement at high schools, although the program does appear to have had a positive effect on graduation rates. These results also suggest that rural schools saw fewer benefits from the program than urban schools, while the effects for charter schools were similar to the effects for traditional public schools.

I. INTRODUCTION

Since the release of *A Nation at Risk* almost thirty years ago, the conspicuous failure of successive waves of school interventions has highlighted the inherent difficulty of raising student achievement at so-called "failing" schools. The School Improvement Grant Program (SIG), promoted by the Obama Administration as the antidote to this chronic underachievement, is the most recent—and arguably the most ambitious—attempt to administer shock therapy to these schools. Since the 2010–2011 school year, SIG has directed over $3.5 billion to more than 1,300 "persistently lowest-performing" schools across the country, in exchange for the adoption of one of four approved models of school improvement: Transformation, Turnaround, Restart, and Closure.[1] Department of Education guidelines allow states to award districts up to $2 million annually to each qualified SIG school. However, in practice, the funding schools have received through SIG has varied, as has the impact of SIG funds on per pupil spending.

Since 2012, the Obama Administration has claimed that SIG is producing "double-digit increases" in the percentage of students deemed proficient in math and English language arts (ELA). However, this claim is based on a correlation and considers only a narrow slice of school performance, i.e. "proficiency," so the true effects of the program remain poorly understood. To date, only one study has rigorously examined the effects of SIG on student achievement. Using school-level data from California, Dee (2012) estimates that SIG had significant positive effects on a school's Academic Performance Index, or API score—a composite measure of school performance calculated annually by the California Department of Education—in its first year of implementation, with the bulk of the gains concentrated in schools that chose to implement the Turnaround model of improvement.

Unfortunately, decades of research on school improvement suggest that the success of even a well-designed intervention is highly dependent on the context in which it is implemented. Thus, since SIG grants have been awarded in all 50 states and the District of Columbia, it is far from certain whether Dee's results are representative of the program as a whole. This is particularly true given that what we do know about SIG suggests that its implementation varies considerably by state. For example, in some states grants were awarded through a competitive process, while in others nearly every eligible school received a grant. Similarly, while in some states schools chose to implement several of the approved improvement models, in

[1] The requirements of these models are defined by federal regulation: Transformation requires that a school incorporate student achievement into teacher evaluations and that the principal be replaced. Turnaround is similar to Transformation, but includes the additional requirement that at least 50 percent of the teaching staff be replaced. Restart requires that a school be closed and reopened as a charter school. Finally, Closure provides minimal funding to assist with the permanent closure of a school.

others virtually all schools chose the Transformation model, which most observers consider to be the least intrusive of the four options (Hurlburt et al. 2011).

This variability in implementation, coupled with the inherent differences that exist between schools and communities, provides ample grounds for interpreting Dee's results narrowly. Accordingly, the primary purpose of this paper is to build upon Dee's work in California by examining the effects of the SIG program in a different context. To that end, I use school-level assessment and graduation data from Texas to test whether receiving a School Improvement Grant leads to higher graduation, completion, or dropout rates and/or increased student achievement, as measured by standardized tests in reading, math, science, social studies, and writing. Further tests are performed to assess the effects of the program for urban versus rural schools and charters versus non-charters. The results of my analysis suggest that in its first year of implementation, SIG has a negative effect on student achievement at elementary and middle schools across almost all subjects and subgroups, and little effect on achievement at high schools, although the program does appear to have a positive effect on graduation rates. Rural schools appear to receive fewer benefits from the program than urban schools, while the effects for charter schools are similar to the effects for traditional public schools—an important result, given the relatively poor performance of these charters.

II. DATA

The data for this paper are drawn from the Common Core of Data and the Texas Department of Education, which provides access to the Texas Assessment of Knowledge and Skills (TAKS) for every campus in the state that reports such data. Prior to the 2011–2012 school year (when Texas switched tests), TAKS data exist for up to five core subjects depending on the grade level, including reading (grades 3–11), math (3–11), science (5, 8, 10, and 11), social studies (8, 10, and 11), and writing (4 and 7).[2] In order to provide a robust evaluation of the impact of SIG, this study uses data for all five subjects, all three proficiency measures, and nine possible subgroups: males, females, Caucasians, African Americans, Hispanic students, economically disadvantaged, at-risk, special education students, and English Language Learners.

The analytical sample for this study consists of all primary and secondary schools in Texas that reported graduation, completion, or dropout rates, and/or TAKS reading, math, science, social studies, or writing assessment data, between the 2007–2008 and 2010–2011

[2] There are no data for the 12th grade because there is no separate test for this grade. Texas students cannot graduate from high schools unless they pass "exit-level" TAKS tests in reading, social studies, math, and science. Consequently, during their junior and senior years of high school, students are given five chances to pass these tests.

school years—approximately 7,800 schools. From these schools, the Texas Department of Education identified those "persistently lowest-achieving" schools that were eligible for SIG funding, as required by federal regulation. More specifically, from a pool of approximately 3,500 Title I-eligible schools in improvement, corrective action, or restructuring, the Department identified approximately 180 schools (roughly five percent) as "persistently lowest-achieving" in Cycle I of SIG. This pool included "Tier I" schools (drawn from the population of schools that received Title I funding), "Tier II" schools (drawn from the population of schools that were eligible for, but did not receive, Title I funds), and a number of "Tier III" schools (other low-performing schools not eligible for Title I funding, but eligible for SIG funding) which received lower priority than Tier I and Tier II schools.

The Tier I and Tier II schools identified as SIG-eligible included schools deemed "persistently lowest achieving" based on schools' average math and reading test scores and their lack of progress in these subjects over the previous two years.[3] Also deemed SIG-eligible were any high schools with graduation rates below 60 percent. Of the Tier I and Tier II schools identified by these criteria, 48 received a SIG award in Cycle I, as did 17 Tier III schools, meaning that graduation and assessment data from these schools in 2010-2011 reflect the impact of the program in its first year of implementation. Of these 65 schools, 53 were high schools and 20 were charter schools. Finally, 63 of the schools receiving grants chose to implement the Transformation model of school improvement, while two schools chose the Turnaround model.

III. METHODOLOGY

The goal of this paper is to estimate the impact of SIG on student achievement during its first year of implementation (2010–2011). However, because there are year-to-year changes in a school's performance that were not attributable to the impact of SIG, simply comparing the test scores and graduation rates of SIG schools before and after implementation will not provide reliable estimates of the program's effect. Econometric techniques can control for the effects of confounding variables such as race, gender, and socioeconomic status, as well as any broader trends in graduation and student achievement that may affect all Texas schools, regardless of their SIG status. In his study of the effects of SIG on student achievement in California, Dee (2012) addresses these issues by using a regression discontinuity model to estimate the effect of SIG eligibility on school performance at various eligibility thresholds. However, in this paper I rely upon the following difference-in-differences model:

[3] Due to the ambiguous language in the regulations governing SIG, different states developed different definitions of "lack of progress" to identify eligible schools. According to Hurlburt et al. (2011), eleven states used a student-level growth measure to determine whether a school had made progress, while the remaining 39 states (including Texas) focused on school-level improvement over time.

$$Achievement_{ct} = \beta_0 + \beta_1 SIG + \beta_2 post + \beta_3 post \times SIG + \beta_4 X + \varepsilon$$

In the above equation, *Achievement* is the academic outcome of interest at campus *c* in year *t*, which may be either the graduation, completion, or dropout rate; the average test score in a given grade and subject; or the percentage of students who met the statewide proficiency or commended standard for a given grade and subject.[4] *SIG* is a dummy variable indicating whether or not a school is part of the first SIG cohort, and *post* is a dummy variable that is equal to "0" for the time period prior to implementation and "1" for the year in which SIG was implemented. The coefficient on the $post \times SIG$ variable represents the estimated increase or decrease in a given achievement measure that is expected at SIG schools, once the expected differences between SIG and non-SIG schools, captured by *SIG* and the statewide trend in proficiency rates captured by *post*, are taken into account.

The full model includes grade-level controls for race, gender, economically disadvantaged status, special education status, English Language Learner (ELL) status, as well as the proportion of students who meet the Texas definition of at-risk youth. For each grade and subject included in the analysis, these controls were generated by dividing the number of tested students in that subgroup by the total number of tested students to find the proportion of tested students in that subgroup for that grade and subject. Additionally, the full version of the model includes both school and year fixed effects, which effectively control for any unobservable campus- or time-invariant characteristics.

A "Title I" dummy variable was constructed by combining the six categories of Title I eligibility from the Common Core of Data into two categories, which was then used to restrict the sample to test the robustness of the results. Similarly, Common Core charter status was reduced to two categories, and the twelve location codes used in the Common Core of Data were combined into two (urban and rural). The rural and charter dummies were then used to restrict the sample so the effects of SIG on urban versus rural schools and charters versus non-charters could be estimated separately.

For each regression, the data are weighted to reflect the number of students represented by a given school for a given performance measure. Thus, for the specifications used to estimate the effect of SIG on graduation, completion, and dropout rates, the data are weighted by school enrollment. For the specifications used to estimate the effect of SIG on reading, math, science, social studies, and writing achievement, the data are weighted by the number of students taking the test in a given grade and subject. Similarly, for each regression performed on a grade-subject pair the standard errors are clustered by campus to account for the possibility that the errors might be

[4] For the purposes of this paper, a "campus" is distinct from a "school," which may include multiple campuses.

correlated within campuses for that grade and subject.[5]

Finally, because the scoring scale for several tests changed between 2008–2009 and 2009–2010, average test scores in all grades and subjects were normalized by subtracting the statewide mean for individuals for a given grade and subject in a given year and dividing the remaining quantity by the standard deviation for that grade, subject, and year. Consequently, while the units for the estimates of SIG's impact on graduation and proficiency rates are percentage points, all estimates for average test scores presented in this paper are expressed in standard deviations.

IV. RESULTS

Table 1 presents summary statistics for the Texas school system and the subpopulation of schools that received SIG awards in 2010–2011. As can be seen from this table, the SIG cohort differs from the broader population of Texas schools in several important ways. Compared to the broader population of schools, SIG schools have lower test scores,[6] lower graduation and completion rates, higher dropout rates, and higher percentages of African American and Hispanic students, as well as more economically disadvantaged, at-risk, special education students, and English Language Learners. Additionally, SIG schools are more urban, more likely to be eligible for Title I funds, and far more likely to be charter and/or high schools than the broader population of schools, meaning that the results for students in grades 9, 10, and 11 are the most important for evaluating the program's overall impact.

For each grade and subject in which a TAKS test was administered, various specifications of the difference-in-differences model were used to test the robustness of the resulting estimates, as illustrated in Table 2, which shows the estimated coefficients for average 10th grade math scores for the "all students" group. In this table, column 1 shows the results for the basic difference-in-differences model without controls; column 2 shows the results including a range of demographic controls; column 3 shows the results for the full model with school and year fixed effects; and column 4 shows the results for the full model when the sample is restricted to schools that were eligible for school-wide Title I funding in 2010–2011. This population of schools bears a greater resemblance to the SIG cohort than the Texas school system as a whole, making it useful for confirming the results from the full sample.

In the most basic version of the model (column 1) the estimated coefficient on SIG is negative and highly significant, suggesting that (prior to receiving an award) SIG schools performed approximately a

[5] An important limitation of this paper arises from the fact that while this method of clustering allows for non-independence within campuses for each grade-subject pair, it does not allow for non-independence within campuses *across* grades and subjects.

[6] Across all tested grades, SIG students scored between .2 and .5 standard deviations below the average Texas student in math, reading, science, social studies, and writing.

Table 1. Summary Statistics

Variable	Texas	SIG
10th Grade Math	0.000	-0.3119
	(0.3788)	(0.3428)
Graduation Rate	87.61	71.75
	(12.01)	(19.76)
Dropout	7.23	18.3
	(8.02)	(13.57)
Completion I	92.77	83.06
	(8.23)	(14.34)
Completion II	93.69	84.78
	(7.35)	(12.78)
Post X SIG	0.0021	0.25
SIG	0.0085	1
Post	0.255	0.25
Percent of tested students who are female	49.51	49.75
	(3.66)	(5.85)
Percent of tested students who are Caucasian	35.89	10.35
	(28.52)	(16.72)
Percent of tested students who are African American	13.91	16.97
	(16.82)	(23.14)
Percent of tested students who are Hispanic	45.68	71.46
	(30.15)	(27.67)
Percent of tested students who are Economically Disadvantaged	53.42	80.68
	(27.13)	(18.32)
Percent of tested students who are "at risk"	41.44	64.61
	(18.28)	(15.97)
Percent of tested students who are Special Ed	6.1	7.71
	(3.28)	(4.86)
Percent of tested students who are Limited English Proficient	9.47	11.05
	(11.84)	(9.78)
High School	0.2427	0.8654
Charter	0.063	0.3077
Rural	0.442	0.2308
Title I eligible	0.7832	0.93
N	7,779	260

Notes: This table shows weighted averages for the Texas school system and the SIG cohort. Standard deviations are in parentheses.

Table 2. Effect of SIG Treatment on 10th Grade Math Scores				
Variable	**(1)**	**(2)**	**(3)**	**(4)**
postXSIG	0.036	0.0274	0.0336	0.0309
	(0.0313)	(0.0321)	(0.0331)	(0.0351)
SIG	-0.3294**	0.0184		
	(0.0606)	(0.0298)		
post	-0.0019	-0.0293**		
	(0.0050)	(0.0056)		
Percent Female		0.0027**	0.0001	0.0001
		(0.0009)	(0.0005)	(0.0006)
Percent Black		-0.0022**	-0.0075**	-0.0067**
		(0.0004)	(0.0009)	(0.0011)
Percent Hispanic		0.0011**	-0.004**	-0.0036**
		(0.0003)	(0.0006)	(0.0007)
Percent Economically Disadvantaged		-0.0037**	-0.0007	-0.0004
		(0.0005)	(0.0004)	(0.0005)
Percent At Risk		-0.0125**	-0.0041**	-0.0039**
		(0.0005)	(0.0004)	(0.0004)
Percent Special Education		-0.0169**	-0.0085**	-0.0077**
		(0.0012)	(0.0008)	(0.0010)
Percent Limited English Proficient		0.004**	-0.0076**	-0.0078**
		(0.0010)	(-0.0013)	(0.0015)
Title I Eligible		-0.0538**	-0.194**	
		(0.0122)	(0.0080)	
School Fixed Effects	No	No	Yes	Yes
Year Fixed Effects	No	No	Yes	Yes
N	6573	6573	6573	4929

Notes: This table shows estimated coefficients for average 10th grade math scores under various specifications. Column 1 shows estimates from the basic model, with no controls included. Column 2 shows estimates including various demographic controls. Column 3 shows estimates including school and year fixed effects. Column 4 limits the sample to schools that are eligible for school-wide Title I programs. All estimates have been normalized and are expressed in standard deviations. Regressions are weighted by the number of students taking the exam at a school. Standard errors that allow for clustering at the school level are in parentheses. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

third of a standard deviation worse than non-SIG schools in 10th grade math, as one might expect given that the program is intended to target only the most "persistently lowest-achieving" schools in the state. In more developed versions of the model (columns 2, 3, and 4) the coefficients on the various demographic controls suggest that schools with a greater percentage of male, African American, Hispanic, economically disadvantaged, at-risk, special education, and ELL students performed less well as a result of these differences. Indeed, as the positive coefficient on SIG in column 2 demonstrates, in the case of 10th grade math, the difference between the performance of SIG and non-SIG schools prior to the implementation of the program is entirely explained by these demographic factors—an important result, since it calls into question one of the underlying assumptions of SIG, that the poor academic performance of grant recipients is at least partly explained by the quality of the education they provide. While this result does not hold for every grade and subject, on average there appears to be little difference between the performance of SIG and non-SIG schools, once demographic factors are taken into account.

Most important for the purposes of this analysis, the estimate on *post×SIG*, despite being relatively stable across all specifications of the model, is not significant in any of them, suggesting that SIG did not have a significant effect on 10th grade math scores. Importantly, there are almost no significant differences between the results for the full sample and those for Title I schools for any grade or subject. Consequently, from this point forward all results presented or discussed are generated using the full sample, unless otherwise indicated. Similarly, from this point forward all results presented are generated using the full version of the model, including demographic controls and school and year fixed effects.

Results for the "All Students" Group

Grade-by-grade estimates of the coefficient on *post×SIG* for reading, math, science, social studies, and writing achievement for the "All Students" group are presented in Tables 3 and 4. For both tables, the coefficients in the "average score" columns represent the expected increase or decrease in the average score for a given subject and grade as a consequence of SIG, expressed in standard deviations. The coefficients in the "percent proficient" and "percent commended" columns represent the expected increase or decrease in the percentage of students who are proficient or commended on a 0–100 scale.

The results for the "All Students" group suggest that the overall impact of SIG on average test scores across grades is mixed, and in many cases the estimated coefficient on *post×SIG* for average test scores is negative and statistically significant. For example, the estimates for average reading scores are negative and statistically significant for grades 4 through 7, with effect sizes approaching one fifth of a standard deviation. Similarly, the estimates for 6th and 8th grade math, 5th and 8th grade science, and 7th grade writing suggest that SIG has had a negative impact on these grades and subjects. Indeed, for no subject in the elementary (3–5) or middle (6–8) grades is the estimated effect on average test scores positive and significant, although the estimates for 3rd grade reading and math are

Table 3. Effects of SIG Treatment on Reading and Math Achievement						
Grade	Reading			Math		
	Average Score	Percent Proficient	Percent Commended	Average Score	Percent Proficient	Percent Commended
3	0.017	1.14	-1.05	0.153	8.81*	3.49
	(0.128)	(2.79)	(7.06)	(0.164)	(3.75)	(6.93)
4	-0.091**	-0.86	-5.63**	-0.009	-1.69	0.97
	(0.034)	(3.13)	(1.58)	(0.041)	(4.14)	(0.84)
5	-0.159**	-1.91*	-7.62	-0.11	-3.9*	-5.63
	(0.023)	(0.94)	(3.05)	(0.078)	(1.70)	(3.57)
6	-0.192**	-5.48**	-9.11**	-0.095**	-3.21*	-2.29*
	(0.037)	(0.88)	(0.88)	(0.011)	(1.33)	(0.95)
7	-0.177**	-3.52**	-3.88**	-0.08	-5.98	0.35
	(0.027)	(1.11)	(0.95)	(0.069)	(4.29)	(0.98)
8	-0.086	-2.05	-3.24	-0.186**	-6.42	-6.35**
	(0.051)	(1.61)	(1.86)	(0.049)	(3.49)	(1.29)
9	-0.013	0.14	-1.51	0.07*	4.33**	1.25
	(0.024)	(0.73)	(0.87)	(0.033)	(1.65)	(1.10)
10	0.047*	2.06*	1.7*	0.034	2.42	0.65
	(0.022)	(0.91)	(0.81)	(0.033)	(1.51)	(0.80)
11	0.009	2.67**	1.34	0.033	3.87**	0.67
	(0.027)	(0.83)	(1.07)	(0.031)	(1.39)	(1.00)

Notes: This table shows estimates of the effects of the SIG treatment on Reading and Math achievement for the full sample, including the controls from column 3 of Table 2 and school and year fixed effects. Estimates of effects on average scores have been normalized and are expressed as standard deviations. For all other estimates the units are percentage points. Regressions are weighted by the number of students taking the exam at a school. Standard errors that allow for clustering at the school level are in parentheses. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

sufficiently large that we cannot rule out a positive effect.

Fortunately, these troubling results do not carry over to the high-school grades (9–11), which, as previously mentioned, are more consequential, since they represent a greater population of schools and students. In general, the estimates for average test scores for these grades suggest that SIG did not have a big effect on high school achievement, although the positive signs on most of these coefficients suggest that any effect was most likely positive.[7] In two cases (10th

[7] Because these results are based on many separate regressions, interpreting the standard errors is problematic without correcting for the number of hypotheses tested, and since there are a total of 27 grade-subject pairs for the All Students group, it could be argued that there are a total of 27 separate hypotheses. Applying a Bonferroni adjustment to the p-values for these hypotheses yields adjusted p-values of .00185 at the 5% significance level and .00037 at the 1% significance level. At these values, many of the estimated effects on average scores for the All

grade reading and 9th grade math), the estimated coefficient is positive and statistically significant; however, in the case of 10th grade reading, the estimate slips below the threshold for significance when the sample is restricted to Title I schools—the only instance in which this restriction makes a significant difference.

Across all grades and subjects, there is suggestive evidence that the effects of SIG were concentrated at the threshold for proficiency, meaning there was a greater increase in the proficiency rate than might be expected, given the increase in test scores. For example, the estimated effect for 3rd grade math proficiency is 8.81 percentage points, despite the fact that the coefficient on average scores is insignificant. Similarly, for 11th grade reading, math, and social studies, and 10th and 11th grade science, the estimated coefficient on proficiency is positive and significant, despite the fact that the coefficient on average scores is not.

One possible explanation for this pattern is that SIG grantees placed additional emphasis on proficiency by focusing on students who are just below the proficiency threshold (so called "bubble students"), or through other means. However, without access to student-level data it is difficult to know for sure. Similarly, it is difficult to evaluate the estimates for the number of students who were "commended," but these results should be interpreted with caution, since the number of students represented by these estimates is small.

Effects of Urbanicity and Charter Status

Table 5 shows the estimated coefficients on *post×SIG* for urban versus rural schools for the "all students" group, which were generated by restricting the sample to each of these subpopulations of schools, using the full model. As can be seen from the table, the results for urban schools appear to be marginally more positive than the results for rural schools. In particular, although only the estimate for 10th grade reading is significant at conventional levels, in virtually every other grade and subject at the high school level, the sign of the coefficients is positive and the estimates are approaching the threshold for significance. Moreover, in several cases the magnitude of the estimated effect is relatively large, approaching a tenth of a standard deviation.

In contrast, nearly all of the estimates for rural schools are negative, and for several grades and subjects (such as 3rd and 9th grade reading, 3rd and 4th grade math, and 10th grade social studies) the results appear to be worse for rural schools than they are for urban schools, although these differences are not necessarily significant. Similarly, the generally negative signs on the estimated coefficients for rural high schools suggest that these schools did

---

Students group, such as 10th grade reading and 11th grade math, fall below the threshold for significance at the 5% level. However, considering the large number of hypotheses involved, it is likely that this approach to adjusting inference is too conservative. Consequently, standard errors for all regressions are presented without adjustment, with the understanding that the risk of false-positives is significant, if difficult to estimate.

Table 4. Effects of SIG Treatment on Science, Social Studies, and Writing Achievement

Grade	Science			Social Studies			Writing		
	Raw Score	Proficient	Commended	Raw Score	Proficient	Commended	Raw Score	Proficient	Commended
3									
4							-0.011	-1.91	3.39
							(0.047)	(1.40)	(2.97)
5	-0.173*	-5.5	-9.23**						
	(0.077)	(3.64)	(3.02)						
6									
7							-0.066**	-0.88	-4.03
							(0.022)	(0.77)	(3.08)
8	0.184**	6.2*	5.15*	0.039	2.52	0.73			
	(0.040)	(3.03)	(2.49)	(0.046)	(2.22)	(1.54)			
9									
10	0.05	3.23*	-0.54	0.032	1.64	1.42			
	(0.029)	(1.35)	(0.72)	(0.031)	(1.09)	(1.10)			
11	0.023	3.64**	-2.87**	0.01	1.97**	-0.15			
	(0.029)	(1.17)	(1.05)	(0.035)	(0.64)	(1.39)			

Notes: This table shows estimates of the effects of the SIG treatment on science, social studies, and writing achievement for grades 3–11, including all of the controls from column 3 of Table 2. Estimates for average scores have been normalized and are expressed as standard deviations. For all other estimates the units are percentage points. Regressions are weighted by the number of students taking the exam at a school. Standard errors that allow for clustering at the school level are in brackets. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

Table 5. Effects of SIG Treatment at Urban vs. Rural Schools

Grade	Rural					Urban				
	Reading	Math	Science	Social Studies	Writing	Reading	Math	Science	Social Studies	Writing
3	-0.58**	-0.269**				0.05	0.177			
	(0.014)	(0.015)				(0.124)	(0.169)			
4	-0.254**	-0.083**			-0.441**	-0.08*	-0.01			0.009
	(0.012)	(0.016)			(0.014)	(0.035)	(0.042)			(0.045)
5	-0.121**	0.246**	0.045*			-0.156**	-0.12	-0.172*		
	(0.016)	(0.019)	(0.022)			(0.026)	(0.082)	(0.085)		
6	-0.18	-0.138**				-0.197**	-0.09**			
	(0.173)	(0.017)				(0.037)	(0.013)			
7	0.091	0.15			-0.032	-0.168**	-0.083			-0.066**
	(0.132)	(0.234)			(0.074)	(0.022)	(0.077)			(0.018)
8	-2.01*	-0.022	-0.116	-0.023		-0.077	-0.19**	-0.186**	-0.038	
	(0.079)	(0.057)	(0.112)	(0.043)		(0.054)	(0.048)	(0.043)	(0.051)	
9	-0.059*	0.031				-0.001	0.073			
	(0.028)	(0.033)				(0.029)	(0.039)			
10	-0.055	-0.022	0.006	-0.101*		0.068**	0.043	0.055	0.055	
	(0.050)	(0.065)	(0.065)	(0.051)		(0.021)	(0.038)	(0.033)	(0.035)	
11	-0.035	0.005	-0.006	-0.044		0.03	0.035	0.019	0.021	
	(0.051)	(0.041)	(0.062)	(0.046)		(0.031)	(0.037)	(0.033)	(0.042)	

Notes: This table shows estimates of the effects of the SIG treatment on achievement at rural vs. urban schools, including all of the controls from column 3 of Table 2 and school and year fixed effects. All estimates have been normalized and are expressed as standard deviations. Regressions are weighted by the number of students taking the exam at a school. Standard errors that allow for clustering at the school level are in brackets. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

not benefit from the program as the urban high schools did—a plausible result, given the criticisms that have been leveled at the program by rural policymakers.[8]

Table 6 shows the estimated coefficients on *post×SIG* for charters versus non-charters for the "all students" group, which were generated by restricting the sample to each of these subpopulations of schools. As can be seen from this table, the relative performance of charters versus non-charters varies by grade. For example, the results for charters appear to be worse than the results for non-charters in 3rd and 5th grade math, but better in 4th grade writing and 6th and 8th grade math, although many of these differences are not significant at conventional levels.

Graduation, Completion, and Dropout Rates

Table 7 presents estimates of the effect of SIG on graduation rates for the full sample, as well as for urban versus rural schools, charters versus non-charters, and the various demographic subgroups. As can be seen from this table, the results for graduation are more encouraging than any of the results discussed so far. In particular, the estimate for the "all students" group suggests that SIG raised overall graduation rates by approximately five percentage points. This result appears to be driven by even greater improvements for African Americans, at-risk students, and special education students. Perhaps the most striking result for graduation is the estimate for African American charter school students, which suggests an increase of nearly 20 percentage points as a result of SIG. Notwithstanding this result, in general the differences between charters and non-charters appear to be modest, as do the differences between urban and rural schools, although the standard errors on many of these estimates are large enough that there may be differences between these groups that are not reflected in such a small sample.

In addition to the results for graduation, Table 8 presents estimates for the effect of the SIG treatment on dropout rates, as well as both of the completion rates tracked by the Texas Education Agency (TEA). According to the TEA, the numerator for completion I consists of students who have graduated or continued in high school, while the numerator for completion II consists of students who have graduated, continued in high school, or received General Education Development (GED) certificates. Interestingly, the estimated effect of SIG on graduation is larger than the estimated effect on completion I, suggesting that some of the increase in graduation rates attributable to SIG may have resulted from the conversion of continuing students (rather than dropouts) into graduates. Similarly, the estimated effect of SIG on completion I is larger than the estimated effect

[8] In a recent survey of state and school-level officials, for example, Scott et al. (2012) found that several SIG requirements, such as the criteria for identifying and funding schools, and the staff replacement requirements of the improvement models, were considered inappropriate for rural schools.

Table 6. Effects of SIG Treatment at Charter vs. Non-Charter Schools

Grade	Charters					Non-Charters				
	Reading	Math	Science	Social Studies	Writing	Reading	Math	Science	Social Studies	Writing
3	-0.186	-0.509*				0.041	0.212			
	(0.131)	(0.229)				(0.144)	(0.162)			
4	-0.173*	0.163			0.466**	-0.105**	-0.013			-0.044
	(0.088)	(0.112)			(0.107)	(0.033)	(0.044)			(0.032)
5	-0.057	-0.217*	0.021			-0.159**	-0.091	-0.174*		
	(0.073)	(0.097)	(0.096)			(0.025)	(0.079)	(0.081)		
6	-0.456**	-0.135				-0.162**	-0.091**			
	(0.110)	(0.114)				(0.016)	(0.010)			
7	-0.204*	-0.137			-0.114	-0.169**	-0.072			-0.062**
	(0.111)	(0.176)			(0.083)	(0.026)	(0.075)			(0.022)
8	0.019	0.131	-0.022	0.079		-0.102	-0.231**	-0.21**	-0.061*	
	(0.068)	(0.069)	(0.103)	(0.187)		(0.059)	(0.012)	(0.027)	(0.025)	
9	-0.09	0.053				-0.01	0.068			
	(0.075)	(0.044)				(0.025)	(0.035)			
10	0.022	0.033	-0.029	0.002		0.046	0.032	0.052	0.031	
	(0.063)	(0.091)	(0.076)	(0.061)		(0.028)	(0.035)	(0.030)	(0.032)	
11	0.069	0.053	0.064	0.113		0.004	0.031	0.021	0.004	
	(0.059)	(0.054)	(0.053)	(0.070)		(0.028)	(0.032)	(0.030)	(0.036)	

Notes: This table shows estimates of the effects of the SIG treatment on average scores at charters versus non-charters, including all the controls from column 3 of Table 2 and school and year fixed effects. All estimates have been normalized and are expressed as standard deviations. Regressions are weighted by the number of students taking the exam at a school. Standard errors that allow for clustering at the school level are in brackets. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

Table 7. Effect of SIG Treatment on Graduation Rates					
Subgroup	All Schools	Urban	Rural	Charters	Non-Charters
All Students	5.17**	4.65*	5.49	6.82	4.79**
	(1.63)	(1.86)	(3.41)	(6.85)	(1.59)
Male	4.95**	4.72*	3.57	5.28	4.78**
	(1.87)	(2.14)	(4.17)	(6.74)	(1.92)
Female	5.18**	4.32*	7.23*	7.68	4.63**
	(1.75)	(1.99)	(3.51)	(8.25)	(1.62)
Caucasian	3.61	5.99*	-0.07	-1.35	3.69
	(2.71)	(2.47)	(2.29)	(11.68)	(2.73)
African American	8.53**	7.93*	6.09	19.87	7.92*
	(3.76)	(3.75)	(9.40)	(8.34)*	(3.72)
Hispanic	2.6	1.71	5.24	6.32	2.15
	(1.70)	(1.85)	(3.52)	(7.57)	(1.63)
Economically Disadvantaged	4.12*	3.43	6.13*	6.02	3.76*
	(1.62)	(1.82)	(2.73)	(8.07)	(1.54)
At Risk	7.39**	8.05**	2.17	3.45	7.35**
	(2.35)	(2.47)	(6.02)	(8.45)	(2.43)
Special Ed	7.98**	6.49	9.67	4.88	7.95*
	(3.07)	(3.32)	(8.65)	(11.70)	(3.22)
ELL	0.06	-0.27	-4.35	-13.05	0.8
	(4.93)	(5.30)	(8.98)	(22.08)	(5.01)

Notes: This table shows estimates of the effect of the SIG treatment on graduation rates, including demographic controls and school and year fixed effects. Regressions are weighted by the number of students at a school. Standard errors that allow for clustering at the school level are in brackets. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

on completion II, suggesting that as a result of SIG, fewer students opted for a GED. Additionally, there is some evidence that the effects of SIG on completion are more modest for rural schools than for urban schools, although the magnitude of the standard errors means we cannot rule out the possibility that there is no difference between the two groups.[9]

V. CONCLUSION

The results of this analysis suggest that in its first year of implementation, SIG had a negative effect on student achievement at elementary and middle schools and little effect on high school achievement, although it does appear to have had a positive effect on graduation rates. These results also suggest that rural schools saw fewer

[9] With no data on completion or dropout rates for the demographic subgroups, it was not possible to estimate the effect of SIG on dropout or completion rates for these groups.

Table 8. Effect of SIG Treatment on Dropout and Completion Rates					
Variable	All Schools	Urban	Rural	Charters	Non-Charters
Dropout	-4.65**	-3.88**	-6.67	-4.93	-4.49**
	(1.35)	(1.09)	(5.20)	(2.54)	(1.43)
Completion 1	2.87*	3.45*	-0.49	5.77	2.55
	(1.43)	(1.60)	(2.75)	(4.09)	(1.46)
Completion 2	2.58	3.01	-0.12	3.63	2.41
	(1.37)	(1.54)	(2.82)	(2.58)	(1.44)

Notes: This table shows estimates of the effect of the SIG treatment on dropout and completion rates, including demographic controls and school and year fixed effects. Regressions are weighted by the number of students at a school. Standard errors that allow for clustering at the school level are in brackets. A single asterisk denotes significance at the 5% level. A double asterisk denotes significance at the 1% level.

benefits from the program than urban schools, while the effects for charter schools were similar to the effects for traditional public schools.

We should be cautious in interpreting these results, for a number of reasons. First, because this analysis was limited to the first year of implementation, these findings must be considered preliminary and subject to revision. Second, because this analysis was limited to school-level data, it was not possible to control for the effects of attrition, which may be significant given the number of youth attending SIG schools who are marginally attached to the education system. Third, because of limitations in the Common Core finance data, it was not possible to fully control for the effects of school spending, which could bias the results. Finally, because schools eligible for SIG are likely to have performed badly in the year prior to receiving the grant, it is possible that any increase in test scores that occurred after the program was implemented reflected a regression to the mean and not the impact of the program itself.[10]

Despite these limitations, the results of this analysis as presented are plausible, if somewhat discouraging, given the scale of federal investment. It should not be surprising to see mixed results, given the generally pessimistic tone of the literature on school improvement. Given what we know about the effects of principal tenure on student achievement, it seems likely that the leadership transitions that occurred at SIG schools in the first year of implementation negatively impacted academic outcomes, meaning that data from subsequent years may paint a more accurate, and potentially favorable, picture of the program's direct impact.[11] In the case of rural

[10] This is one possible explanation for the relatively positive results for third grade, which would probably be the first to reflect the enrollment of a more academically capable group of students.

[11] Béteille et al. (2012) found that Miami schools with first-year principals had lower achievement gains than other schools. Similarly, Miller (2009) found that schools in North Carolina experienced a decline in student achievement

schools, these leadership transitions were probably particularly rough, given the difficulty of attracting qualified principals to rural areas. However, without additional information on district hiring practices, it is difficult to say how important this factor was. Similarly, while it is likely that certain demographic subgroups were targeted for improvement as part of the turnaround process, without additional information on how this demographic targeting occurred, any attempt to account for it would be speculative. In particular, it is difficult to know whether the absence of a positive effect for a particular group reflects a lack of effort or a lack of success, especially since different schools likely took different approaches to raising achievement.

Texas is unique in that it decided to award a large percentage of its SIG grants to low-performing charter schools—a confusing policy, since one of the primary motivations for encouraging the growth of charters is to introduce a measure of market discipline into the education sector.[12] Since charter grantees are failing to make significant progress, despite receiving additional funding from SIG, it seems reasonable to ask why nearly all of these schools are implementing Transformation as opposed to Closure. The answer, of course, is that in many states, low-performing charters (like low-performing district schools) are rarely closed (Stuit 2010). Nationwide, only two percent of the first SIG cohort chose Closure over the other three models (Hurlburt et al.). Thus, the real debate going forward may have less to do with the merits of charters versus non-charters than with the merits of school closure versus school turnaround generally.

Arguably, the results of this analysis bolster the case for closure, since SIG schools for the most part failed to make progress, despite receiving additional funding and support. However, since Texas used absolute performance (rather than some measure of school value added) to identify which schools were eligible for a SIG grant, and since essentially all of the difference in performance between SIG and non-SIG schools can be explained by demographics, it could also be argued that these schools are not really "failing" in the first place. By assuming that poor absolute performance reflects poor teaching and/or school management, it is possible that the education officials responsible for implementing SIG are repeating the mistakes of No Child Left Behind by identifying the wrong schools for improvement. If this is the case, we should not be surprised that replacing the leadership at these schools is not leading to better academic outcomes.

---

in the years immediately following a change in leadership.

[12] Interestingly, including the dummy variables *charter* and *SIG×charter* in various alternative specifications of the model yields negative and statistically-significant estimates for both coefficients across most grades and subjects, implying that not only are charters in Texas performing poorly relative to the rest of Texas schools, but that grant-receiving charters are an unusually low-performing bunch, even after their charter status, SIG status, and demographic characteristics are taken into account.

Before we pass judgment on the merits of SIG, we should remember that implementation of the program has varied considerably by state, and consequently, the impacts of the program may also have varied. While the overwhelming majority of Texas schools chose to implement the Transformation model, the two schools that implemented the Turnaround model saw generally positive results in their first year.[13] Thus, since Dee (2012) found that the positive effects of SIG in California were concentrated in Turnaround schools, it is possible that the results of the two studies may prove consistent with one another, insofar as they reflect the differential impacts of the two models.

[13] Specifically, Azleway Charter School Pine Mountain saw a 17-point increase in 9th grade math proficiency, while Floresville Choice program saw a 24-point increase in 10th grade ELA proficiency and a 6-point decrease in 10th grade math proficiency.

VI. REFERENCES

Béteille, T., D. Kalogrides, and S. Loeb. 2012. "Stepping stones: Principal career paths and school outcomes." *Social Science Research* 41, no. 4: 904-919.

Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin. 2012. "Estimating the Effect of Leaders on Public sector Productivity: The Case of School Principals." NBER Working Paper 17803, National Bureau of Economic Research, Cambridge, MA.

Dee, Thomas. 2012. "School Turnarounds: Evidence from the 2009 Stimulus." NBER Working Paper 17990, National Bureau of Economic Research, Cambridge, MA.

Hurlburt, S., K. C. Le Floch, S. B. Therriault, and S. Cole. 2011. "Baseline Analyses of SIG Applications and SIG-Eligible and SIG-Awarded Schools." NCEE 2011-4019, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

Miller, Ashley. 2009. "Principal turnover, student achievement and teacher retention," Unpublished Manuscript, Princeton University, Princeton, NJ.

Rhim, L. M., Kowal, J. M., Hassel, B. C., & Hassel, E. A. 2008. School turnarounds: A review of the cross-sector evidence on dramatic organizational improvement," Lincoln, IL: Public Impact, Academic Development Institute.

Scott, Caitlin, Jennifer McMurrer, Shelby McIntosh, and Kenne Dibner. 2012. *Opportunities and Obstacles: Implementing Stimulus-Funded School Improvement Grants in Maryland, Michigan, and Idaho.* Washington, DC: Center on Education Policy.