Effects of Delivery Model and Instructional Hours on Elementary Science Proficiency

Prepared for Arlington Public Schools

November 2013



In this report, Hanover Research investigates how changing the mode of instructional delivery and the number of instructional contact hours may affect student outcomes on third and fifth grade standardized tests in science.

HANOVER RESEARCH

TABLE OF CONTENTS

EXECUTIVE SUMMARY & KEY FINDINGS

INTRODUCTION

In this report, Hanover Research analyzes Arlington Public Schools' (APS) third and fifth grade student performance on the Virginia Standards of Learning (SOL) exam in science.[1] We use a linear regression model to measure the effect of various instructional delivery models and average instructional contact hours on student outcomes. Using data on third and fifth grade students, we examine scale scores, pass/fail status, and proficiency ratings.

KEY FINDINGS

- **Among third graders**, none of the coefficients on the variables of interest are significant using a 95 percent confidence interval. This implies that, in our sample of third grade students, **the effects of both instructional model and average hours of instruction are not statistically different from zero after controlling for other factors**. Notably, we observe significant coefficients on most of our demographic control variables, which suggests that **variation in instructional models and in average hours of instruction cannot explain third grade outcomes as well as variation among students themselves**.
- Across all three models of fifth-grade students, **Average Hours of Instruction is both positive and statistically significant** using a 99 percent confidence interval. This is **strong evidence that fifth-grade students who have additional instruction hours in science can be expected to earn higher scores on the SOL test** and thus to have higher probabilities both of passing and of passing at an advanced level.
- **"Classroom Teacher" seems to be the best instructional model for fifth grade students.** For fifth graders, with scale score as the outcome variable, each instructional model has a lower outcome score compared to classroom teacher (Instructional Model One). These results suggest that other types of instructional delivery may be correlated with worse outcomes on the SOL tests in science. For third graders, no single instructional model appears to be superior (or inferior) than the other models.
- Demographic characteristics are correlated with SOL outcomes: For both grades, students with LEP status, economically disadvantaged students, students with SPED status, and black and Hispanic students have lower SOL science score outcomes than their comparison groups[2].

[1] More information about the SOL exam in science can be found at http://www.doe.virginia.gov/testing/sol/standards_docs/science/

[2] The comparison groups for all cases are usually the students who are not in that group, for example LEP's comparison group is "Non LEP students". One exception to this is the race categories where the students are compared to those students who are categorized as white.

- Students in the immersion program have lower SOL science outcomes than students not in the immersion program.

Table I, below, summarizes the main findings of our study.

Table I: Summary of Report Findings

Variable Name	Grade 3	Grade 5
Type of Instructional Model	No relationship	"Classroom Teacher" produces better outcomes than other instructional models.
Average Hours of Instruction	No relationship	An additional hour of instruction is correlated with higher SOL Science outcomes.
Demographics – Gender	Female students have slightly lower SOL outcomes	Female students have lower SOL outcomes in two out of three measures
Demographics – Race	Black and Hispanic students have lower outcomes compared to white students.	Black and Hispanic students have lower outcomes compared to white students.
Special Status – SPED, LEP, Economically Disadvantaged	Associated with lower SOL science outcomes.	Associated with lower SOL science outcomes.
Immersion School	Associated with lower SOL science outcomes.	Associated with lower SOL science outcomes.

SECTION I: DATA & METHODOLOGY

DATA

Arlington Public Schools provided Hanover Research with data on student SOL scores and additional variables for 3,242 students during the 2012-13 school year, representing a total of twenty-two schools. Each student in the dataset is uniquely identified by his or her SIR number (i.e., there are no duplicates). Of these students, 1,628 are in Grade 3 and 1,613 are in Grade 5. In addition, there was one student in Grade 4, but since we are only interested in the SOL scores of third and fifth graders, we drop this student from the dataset before performing our analysis.

We examine three student outcome measures related to Standards of Learning in science.

- Scale scores on the Standards of Learning test in science
- A binary variable indicating whether a student passed the SOL science test, as opposed to failing the SOL science test (i.e. a score above or below 400)
- A binary variable indicating whether a person passed/advanced the SOL science test as opposed to either passed with only a proficient score or failed to pass (i.e. a score above or below 500)

Figure 1.1 shows how these measures are related, with the pass/fail and proficiency indicators using the scale score as their base.

Figure 1.1: Science SOL Scale Scores, Pass/Fail Indicators, and Proficiency Ratings[3]

	Scale score	Pass/Fail Indicator	Proficiency Rating
SCORE RANGE (LOW)	0 – 399	Fail	Fail
SCORE RANGE (MID)	400 – 499	Pass	Proficient
SCORE RANGE (HIGH)	500 – 600	Pass	Advanced

In our regression model, these three outcome measures serve as our dependent variables, whose values we predict using data on related, explanatory variables.

Figure 1.2 depicts the distribution of the scale scores for third and fifth grade respectively.

[3] Source: http://www.doe.virginia.gov/administrators/superintendents_memos/1998/inf179.html

Figure 1.2: Distribution of Science SOL Scale Scores

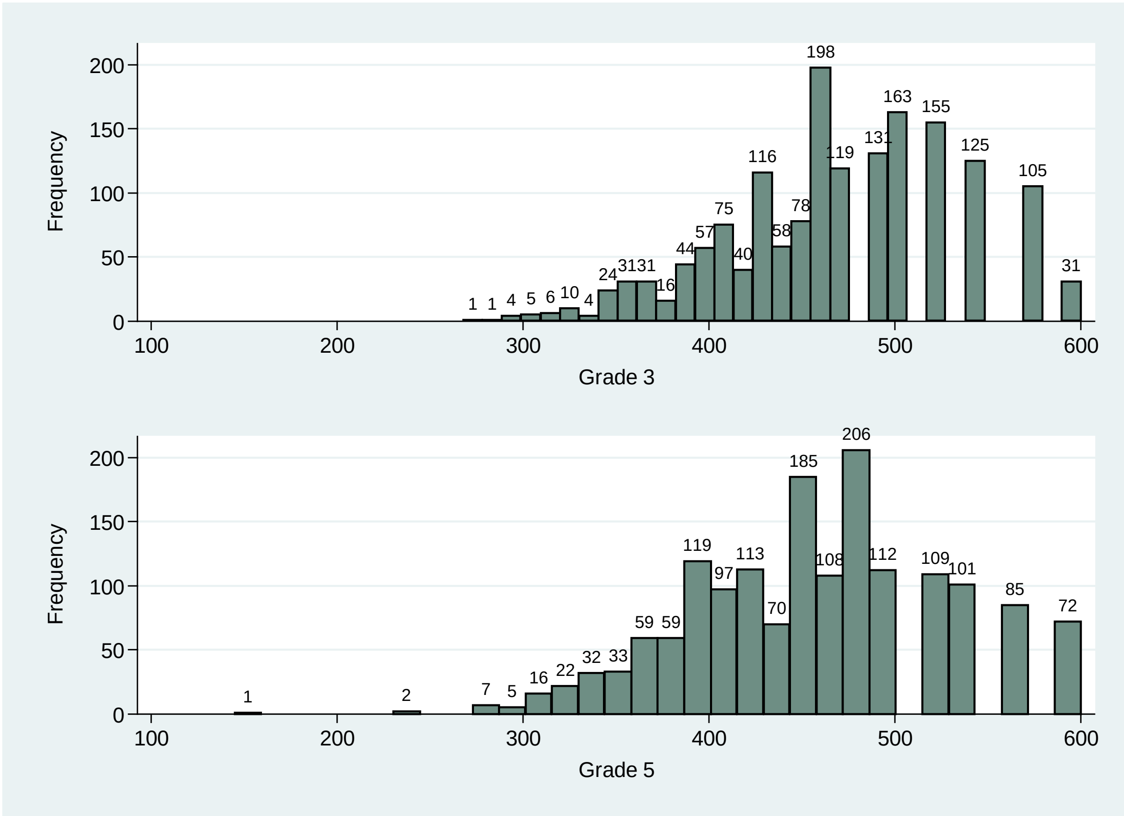

Figure 1.3 describes distribution of race in the student data. We observe that slightly more than half of the students are categorized as white, so they serve as the reference category throughout our analysis.

Figure 1.3: Race Distribution and Description of Numeric Codes

Description	Frequency	Percentage
Asian	272	8.39%
Black	335	10.34%
Hispanic	792	24.44%
White	1,638	50.54%
Other	204	6.29%
Total	**3,241**	**100%**

Other demographic features of students within the dataset include the following.

- 1,650 (or 50.91 percent of the students) are **female**.
- 534 (or 16.48 percent of the students) have a **SPED** designation
- 836 (or 25.79 percent of the students) have an **LEP** designation
- 891 (or 27.49 percent of the students) have an **economically disadvantaged** designation
- 326 (or 10.06 percent of the students) attend an **immersion** school (either Claremont Immersion or Francis Scott Key ES).

Moreover, there are five categories of **instructional delivery models** in the data. Figure 1.4 describes each category and lists the number of students by grade. *Instructional model is one of the predictor variables of primary interest in this study.*

Figure 1.4: Instructional Delivery Models

Instructional Delivery Model	Grade 3		Grade 5	
	Freq.	Pct.	Freq.	Pct.
Classroom Teacher[4] (Model One)	867	53.26%	483	29.94%
Classroom Teacher plus enrichment (Model Two)	374	22.97%	202	12.52%
Rotate teachers for science instruction (Model Three)	215	13.21%	468	29.01%
Rotate teachers for science instruction plus enrichment[5] (Model Four)	0	0%	213	13.21%
Science specialist (Model Five)	172	10.57%	247	15.31%
Total	**1628**	**100.00%**	**1613**	**100.00%**

The other predictor variable of primary interest is **average instruction hours in science**. The dataset contains the frequency of teachers indicating a given number of hours in science instruction that a student receives in a particular school and grade. Instruction hours were coded into categories as outlined in Figure 1.5.

Figure 1.5: Hours of Science Instruction

Grade 3 Variable Name	Grade 5 Variable Name	Description	Average Time (Hours)
Lessthan1Third	Lessthan1Fifth	Science instruction occurred less than one hour per week	0.5 hours

[4] In our analysis, Classroom Teacher (Instructional Model One) serves as the reference category against which the performance of other instructional models are compared.

[5] We do not have records for any third grade students with instructional delivery model 4 (rotating science teacher plus enrichment).

Grade 3 Variable Name	Grade 5 Variable Name	Description	Average Time (Hours)
OneHourThird	OneHourFifth	Science instruction occurred between 1hr-1hr 59 min per week	1.5 hours
TwoHourThird	TwoHourFifth	Science instruction occurred between 2hr-2hr 59 min per week	2.5 hours
ThreeHourThird	ThreeHourFifth	Science instruction occurred between 3hr-3hr 59 min per week	3.5 hours
FourHourThird	FourHourFifth	Science instruction occurred 4 or more hours per week	4 hours

In the table, we include an average time per category, which is estimated as the mid-point of each range. The one exception is the highest category (more than four hours) where we use the minimum of the range (4 hours). Since it is possible for multiple teachers to estimate hours of science instruction for any given student, we use the midpoints of each category and then take a weighted average.

Thus, for example, if three teachers estimate one particular student's instruction hours such that *two* of them estimate Lessthan1Third (0 to 1 hour ➔ midpoint of 0.5 hours) and *one* of them estimates OneHourThird (1 to 1 hour 59 minutes ➔ midpoint of 1.5 hours), then the weighted average for this student will be:

$$\frac{(2 \times 0.5) + (1 \times 1.5)}{3} \approx 0.83 \text{ hours} = 50 \text{ minutes}$$

Finally, we also analyze the relationship between science outcomes and the **average survey response** of students, parents, and teachers. The specific responses examined were collected through a survey that was administered by the district. Among other questions (not related to the district's science instruction), the survey asked parents, students, and teachers about their satisfaction with the district's science program (parents and students) or the appropriateness of the amount of time students spend learning science in school (teachers). Figure 1.6 shows the variable names and the specific survey questions. Numeric responses to these questions on a scale of one to four were aggregated for each school and then averaged to provide one score for each student.

Figure 1.6: Survey Questions

Variable Name	Specific Question	Range
ParentSBSResponse	"Please rate your level of satisfaction with the education your child is receiving for each of the following subject areas."	(1=very dissatisfied—4=very satisfied)
StudentSBSResponse	"Please rate your level of agreement with the statement, 'I enjoy learning about science'"	(1=strongly disagree—4=strongly agree)
TeacherSBSResponse	"Please rate your level of agreement with the statement, 'Students spend enough time learning about science.'"	(1=strongly disagree—4=strongly agree)

METHODOLOGY

Since each outcome variable is based on the same test, it is likely that the same explanatory variables will be significant predictors of all three outcomes. However this result is not guaranteed. Therefore, we analyze them separately by grade and then compare the results.

We specify the equation for scale scores as a linear regression model with robust standard errors. The binary response variables (passing versus non-passing, passing/advanced versus passing/proficient or non-passing) are specified as linear probability models. As a final robustness check, we re-run our models using alternative specifications and include the results in an Appendix. These alternative models employ school-level fixed effects to control for school-wide differences among students. Since average survey response is also a school-level variable, we must exclude it from our alternative specifications to avoid over-fitting the model.

We use the method of Ordinary Least Squares (OLS) to estimate parameters of the following linear equation. Separately by grade, for each student *(i)*, we run a separate model for each outcome variable, SOL Score, a binary variable for student passing or not, and a binary variable for pass/advanced:

$$
\begin{aligned}
(Outcome)_i = \alpha &+ \beta_1(Classroom\ Teacher\ Plus\ Enrichment)_i \\
&+ \beta_2(Rotating\ Science\ Teacher)_i \\
&+ \beta_3(Rotating\ Teacher\ Plus\ Enrichment)_i \\
&+ \beta_4(Science\ Specialist)_i \\
&+ \beta_t(Average\ Instruction\ Hours)_i \\
&+ \delta(Average\ Survey\ Responses)_i \\
&+ \gamma(Demographics)_i + \varepsilon_i.
\end{aligned}
$$

Here, $\beta_1$through β_4 are coefficients on the dummy variables indicating the instructional delivery model, with Classroom Teacher serving as the reference category. We are primarily interested in these four coefficients, along withβ_t, which is the coefficient on average hours of science instruction. The error term, ε_i, is assumed to be random with mean zero and constant non-zero variance.

Interpreting Regression Results

A coefficient estimated by an OLS regression model indicates the amount by which the outcome variable (e.g., SOL scale score) changes in response to a one-unit change in a given predictor variable. A positive coefficient indicates a positive relationship between the two variables. In other words, when a continuous predictor variable increases (or decreases), the outcome variable increases (or decreases). The coefficient estimates the magnitude of the change while holding all other predictor variables constant. In the case of a categorical predictor variable, such as gender, we interpret the coefficient in relation to the designated reference group. For example, a positive coefficient for gender indicates that females earn a higher scale score on average than males.

With linear probability models, we interpret the estimated coefficients differently, based on the binary nature of the outcome variable (e.g., pass/fail status). In contrast to continuous variables, binary variables, by definition, only assume one of two values. In the context of the present analysis, we assign a value of 1 if a student passed (e.g., earned a score of 400 or more) and a value of 0 otherwise. Accordingly, a coefficient in a linear probability model indicates the estimated change in the *probability* that a student will pass following a one-unit change in a given predictor variable (holding all other predictor variables constant). A positive coefficient still indicates a positive relationship—when a continuous predictor variable increases (decreases), the estimated probability increases (decreases). Similarly, we continue to interpret the coefficient of a categorical predictor variable relative to the designated reference group. For instance, a positive coefficient for gender indicates that females are more likely to pass on average than males.

In our analysis of instructional delivery models and instruction hours, positive and significant estimates for any of the coefficients $\beta_1$through β_4 will imply that this particular model of instruction improves the outcome measure significantly more than the reference group's model. A positive and significant coefficient estimate for β_t will imply that the outcome variable increases by the amount of the coefficient estimate, given one additional hour of science instruction. The other independent variables in the final model are used to control for any correlations that might otherwise bias our results. However, each of these coefficients can be interpreted similarly.

SECTION II: RESULTS & INTERPRETATION

GRADE 3

Figure 2.1 displays the estimated coefficients from our regression model for third grade students. The primary variables of interest appear first.

Figure 2.1: Grade 3 Regression Coefficients

Predictor Variables	Outcome Variables (Grade 3)		
	SCALE SCORE	PASS/FAIL[6]	PROFICIENCY RATING[7]
Classroom Teacher Plus Enrichment[8] (Model Two)	2.0254	0.0135	0.0038
Rotating Science Teacher (Model Three)	-4.1626	-0.0025	-0.0471
Science Specialist[9] (Model Five)	-3.5811	-0.0945*	0.0760
Average Hours of Instruction	0.2604	0.0034	-0.0005
Gender (Female)	-6.2568**	0.0013	-0.0376*
Race (Asian)[10]	-10.8198**	0.0187	-0.0981**
Race (Black)	-28.5491***	-0.1107***	-0.2061***
Race (Hispanic)	-22.8163***	-0.0849***	-0.1827***
Race (Other)	-8.0652	-0.0444	-0.1001*
SPED	-43.2983***	-0.2379***	-0.2018***
LEP	-17.1420***	-0.0605*	-0.0968***
Economically Disadvantaged	-38.4002***	-0.1633***	-0.1813***
Average Parent Survey Response	-8.0434	-0.2555***	0.2510***
Average Student Survey Response	3.2274	0.0229	-0.0045
Average Teacher Survey Response	16.1855***	0.1046***	0.0431
Constant	454.8887***	1.3876***	-0.4620
Observations	1,543	1,543	1,543
R-squared	0.3222	0.2214	0.1630

The models were estimated using ordinary least squares with robust standard errors.
*** p<0.01, ** p<0.05, * p<0.1

Contrary to our expectations, none of the coefficients on the variables of interest are significant using a 95 percent confidence interval. This implies that, in our sample of third grade students, the effects of both instructional model and average hours of instruction are not statistically different from zero after controlling for other factors such as demographic characteristics.

[6] This is a linear probability model.
[7] This is a linear probability model.
[8] Classroom Teacher (Instructional Model One) is the reference category.
[9] Rotating Teacher plus Enrichment (Instructional Model Four) is excluded due to lack of observations.
[10] White is the reference category.

Notably, we observe significant coefficients on most of our demographic control variables, which suggests that variation in instructional models and in average hours of instruction cannot explain third grade outcomes as well as variation among students themselves:

- **SPED Status is associated with lower SOL scores.** The estimated coefficients on SPED (indicating special education status) are negative and significant in all three models. In the first model, with scale score as the outcome variable, we expect SPED students to earn roughly 43 fewer points on average than non-SPED students. Moreover, negative coefficients in the second and third models, both with binary outcome variables, imply that SPED students are *less likely* to pass the SOL test than non-SPED students, and they are also *less likely* to pass the SOL test at an advanced level than non-SPED students.
- **Economically Disadvantaged status is associated with lower SOL scores.** If a student is economically disadvantaged, their expected Science SOL score is 38.4 points lower than a non-economically disadvantaged student, and that student is 16.3 percent less likely to pass, and 18.1 percent less likely to pass/advanced. All of these results are statistically significant results.
- **LEP status is associated with lower SOL scores.** If a student is limited English proficient, their expected science SOL score is 17.1 points lower than a non-LEP student, and they are 9.7 percent less likely to pass/advanced—both of which statistically significant results. They are also less likely to pass than a non-LEP student (Model 2), although this relationship is not as strong as it is only significant at the 10% level.
- **Hispanic and black students have lower SOL science outcomes.** Both Hispanic and black students have lower SOL scores, are less likely to pass or pass/advanced compared to white students, and this relationship is statistically significant at the 1 percent level.
- **Teacher response survey.** A one point increase in the average teacher response to the survey question stating the level of agreement with the question “Students spend enough time learning about science” leads to an expected increase of 16.2 points in SOL scale score and a 10.4 percent increase in the probability of passing, both of which are statistically significant results. However, changes in the survey response score do not affect the probability of being pass/advanced.

Grade 5

Figure 2.2 displays the estimated coefficients from our regression model for fifth grade students. The primary variables of interest appear first and have light green backgrounds.

Figure 2.2: Grade 5 Regression Coefficients

Predictor Variables	Outcome Variables (Grade 5)		
	Scale Score	Pass/Fail[11]	Proficiency Rating[12]
Classroom Teacher Plus Enrichment[13] (Model Two)	-14.2676***	0.0232	-0.1593***
Rotating Science Teacher (Model Three)	-19.5039***	-0.0262	-0.1485***
Rotating Teacher Plus Enrichment (Model Four)	-3.3794	0.0279	-0.0633*
Science Specialist (Model Five)	-34.0204***	-0.0279	-0.3126***
Average Hours of Instruction	17.6540***	0.0730***	0.0894***
Gender (Female)	-10.3323***	-0.0094	-0.0673***
Race (Asian)[14]	-18.0267***	-0.0290	-0.2184***
Race (Black)	-50.4745***	-0.2344***	-0.2679***
Race (Hispanic)	-34.3509***	-0.1571***	-0.2280***
Race (Other)	6.5819	-0.0310	0.0055
SPED	-32.2728***	-0.1535***	-0.1356***
LEP	-27.8787***	-0.1705***	-0.0764***
Economically Disadvantaged	-26.6890***	-0.1057***	-0.1110***
Average Parent Survey Response	-5.0421	0.0616	-0.1399
Average Student Survey Response	-7.5684	0.0009	-0.0009
Average Teacher Survey Response	8.8019	0.0253	0.1019***
Constant	464.7132***	0.4273	0.4043
Observations	1,357	1,357	1,357
R-squared	0.4084	0.2616	0.2407

The models were estimated using ordinary least squares with robust standard errors.
*** p<0.01, ** p<0.05, * p<0.1

All three of the regression models for fifth graders have a higher R-squared than those for their third grade counterparts, meaning that they can explain more of the differences in outcomes among fifth graders than among third graders. Again, demographic variables are often significant, but many of our primary variables of interest are significant as well.

[11] This is a linear probability model.
[12] This is a linear probability model
[13] Classroom Teacher (Instructional Model 1) is the reference category.
[14] White is the reference category.

In particular, across all three models, **Average Hours of Instruction** is both positive and statistically significant using a 99 percent confidence interval. This is strong evidence that fifth-grade students who have additional instruction hours in science can be expected to earn higher scores on the SOL test and thus to have higher probabilities both of passing and of passing at an advanced level.

Additionally, we see evidence that the instructional model "classroom teacher" is superior to most other instructional models. Observe that compared to the classroom teacher, students in instructional model "Classroom Teacher Plus Enrichment" are expected to score 14.3 points lower in SOL scale score, students in instructional model "Rotating Science Teacher" are expected to score 19.5 points lower, and students in instructional model "Science Specialist" are expected to score 34.0 points lower in SOL scale scores. All of these results are statistically significant at the 1 percent level. These results suggest that, compared to the reference group Classroom Teacher (Instructional Model One), other types of instructional delivery *may be correlated with worse outcomes* on the SOL tests in science.

In addition, demographic variables continue to be significant determinants:

- **SPED Status is associated with lower SOL scores in 5th grade as well**. The estimated coefficients on SPED (indicating special education status) are negative and significant in all three models. In the first model, with scale score as the outcome variable, we expect SPED students to earn roughly 32 fewer points on average than non-SPED students, and are 15.3 percent less likely to pass, and 13.5 percent less likely to pass/advanced, all statistically significant results.
- **Economically Disadvantaged status is associated with lower SOL scores**. If a student is economically disadvantaged, their expected science SOL score is 26.7 points lower than a non-economically disadvantaged student, and is 10.6 percent less likely to pass, and 11.1 percent less likely to pass/advanced, all statistically significant results.
- **LEP status is associated with lower SOL scores.** If a student is limited English proficient, their expected science SOL score is 27.8 points lower than a non-LEP student, and 17.0 percent less likely to pass and 7.6 percent less likely to pass/advanced, all statistically significant results.
- **Hispanic and black students have lower SOL science outcomes.** Similar to the 3rd grade outcomes, in 5th grade both Hispanic and black students have lower SOL scores, are less likely to pass or pass/advanced compared to white students, and this relationship is statistically significant at the 1 percent level.

EFFECT OF IMMERSION SCHOOLS

Arlington Public Schools requested that we include immersion school status as one of our predictors. There are two schools in the dataset that are immersion schools – Claremont Immersion and Francis Scott Key ES. The immersion variable is perfectly correlated with other school-level variables under consideration (notably, Instructional Model Five – Science

Specialist). Including a dummy variable for immersion status would thus invalidate our regression model, so instead we perform simple t-tests to look for differences in means between immersion and non-immersion schools.

A t-test takes the mean value of a given variable for two groups, whose variance is assumed to be identical, and then subtracts one mean from the other. We begin with the hypothesis that the true, unobservable mean of these two groups is the same, so this difference should equal zero. However, in finite samples, the means of two subgroups are often different, and we would like to know the likelihood of observing a given difference. Using statistical theory, it can be shown that the difference in sample means divided by the scaled sample standard deviation is distributed as a *student's t-distribution*, so we can perform statistical tests after setting a significance level.

In Figure 2.3, a p-value of less than 0.05 is evidence against the starting hypothesis that these two groups are the same. Thus, for Grade 3, a p-value of 0.0005 indicates that the difference in mean scale score between non-immersion and immersion schools (470.69 - 453.43 = 17.26) is statistically different from zero at the 0.05 percent significance level, i.e. *extremely* significant.

We find that for both grades and all three outcome measures, immersion school's SOL outcome are <u>*lower*</u> compared to non-immersion schools.

Figure 2.3: T-test for Difference in Means between Immersion and Non-immersion Schools

	Group	Observations	Mean	P-value
GRADE 3	**Scale Score**			
	Non-immersion	1456	470.69***	0.0005
	Immersion	172	453.43	
	Pass/Fail			
	Non-immersion	1456	88.53%***	0.0002
	Immersion	172	78.49%	
	Proficiency Rating			
	Non-immersion	1456	36.26%*	0.0868
	Immersion	172	29.65%	
GRADE 5	**Scale Score**			
	Non-immersion	1459	460.11***	0.0000
	Immersion	154	427.35	
	Pass/Fail			

	Group	Observations	Mean	P-value
	Non-immersion	1459	81.91%***	0.0000
	Immersion	154	67.53%	
	Proficiency Rating			
	Non-immersion	1459	31.53%***	0.0000
	Immersion	154	12.34%	

*** p<0.01, ** p<0.05, * p<0.1

FUTURE RESEARCH

In future projects, Hanover Research could improve this analysis by including more student-level variables as predictors. For example, in our model, we do not explicitly control for individual unobservable factors, such as student ability or motivation. In order to isolate the effects of a particular program or instructional delivery model, it may help to include such factors, as they are likely to confound the results. One possible proxy for student ability is GPA. We may be interested in segmenting the students in some way, possibly by GPA, and specifying a model to determine if delivery model and instruction hours have different effects on different groups of students.

Appendix: Robustness Checks

Grade 3

Figure A1: Grade 3 Robustness Check Regression Coefficients

Predictor Variables	Outcome Variables (Grade 3)		
	Scale Score	Pass/Fail[15]	Proficiency Rating[16]
Classroom Teacher Plus Enrichment[17] (Instructional Model Two)	4.6510	0.0260	0.1005
Rotating Science Teacher (Instructional Model Three)	-18.6363**	-0.0237	-0.1754**
Science Specialist[18] (Instructional Model Five)	8.3889	-0.0538	0.3895***
Average Hours of Instruction	13.4904	0.0175	0.2166***
Gender (Female)	-6.2670**	0.0014	-0.0360
Race (Asian)[19]	-15.1506***	-0.0032	-0.1218***
Race (Black)	-23.4949***	-0.0948***	-0.1747***
Race (Hispanic)	-17.0862***	-0.0649**	-0.1531***
Race (Other)	-4.1992	-0.0348	-0.0740
SPED	-46.2304***	-0.2481***	-0.2136***
LEP	-9.4489**	-0.0396	-0.0543
Economically Disadvantaged	-34.1760***	-0.1485***	-0.1582***
School Fixed Effects?	Yes	Yes	Yes
Constant	457.2037***	0.9468***	-0.2201
Observations	1,543	1,543	1,543
R-squared	0.3887	0.2670	0.1949

The models were estimated using ordinary least squares with robust standard errors.
*** p<0.01, ** p<0.05, * p<0.1

[15] This is a linear probability model.
[16] This is a linear probability model.
[17] Classroom Teacher (Instructional Model One) is the reference category.
[18] Rotating Teacher plus Enrichment (Instructional Model Four) is excluded due to lack of observations.
[19] White is the reference category.

GRADE 5

Figure A2: Grade 5 Robustness Check Regression Coefficients

PREDICTOR VARIABLES	OUTCOME VARIABLES (GRADE 5)		
	SCALE SCORE	PASS/FAIL[20]	PROFICIENCY RATING[21]
Classroom Teacher Plus Enrichment[22] (Instructional Model Two)	6.7804	0.0376	-0.0188
Rotating Science Teacher (Instructional Model Three)	-15.9496**	-0.0512	-0.0860*
Rotating Teacher Plus Enrichment (Instructional Model Four)	-27.0618***	-0.0717	-0.1572**
Science Specialist (Instructional Model Five)	-31.8142***	-0.0895	-0.2428***
Average Hours of Instruction	35.7631***	0.0759	0.1873***
Gender (Female)	-10.4676***	-0.0114	-0.0672***
Race (Asian)[23]	-14.7019**	-0.0138	-0.2019***
Race (Black)	-48.3232***	-0.2257***	-0.2524***
Race (Hispanic)	-31.9417***	-0.1469***	-0.2171***
Race (Other)	7.3527	-0.0301	0.0123
SPED	-34.3421***	-0.1643***	-0.1425***
LEP	-28.7312***	-0.1802***	-0.0784***
Economically Disadvantaged	-24.1617***	-0.0960**	-0.0991***
School Fixed Effects?	Yes	Yes	Yes
Constant	346.0***	0.601***	-0.261
Observations	1,357	1,357	1,357
R-squared	0.4336	0.2737	0.2572

The models were estimated using ordinary least squares with robust standard errors.
*** p<0.01, ** p<0.05, * p<0.1

[20] This is a linear probability model.
[21] This is a linear probability model
[22] Classroom Teacher (Instructional Model One) is the reference category.
[23] White is the reference category.

PROJECT EVALUATION FORM

Hanover Research is committed to providing a work product that meets or exceeds member expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.

http://www.hanoverresearch.com/evaluation/index.php

CAVEAT

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties which extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every member. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Members requiring such services are advised to consult an appropriate professional.

HANOVER RESEARCH

1750 H Street NW, 2nd Floor
Washington, DC 20006

P 202.756.2971 F 866.808.6585
www.hanoverresearch.com