The Impact of the Disadvantaged Student Supplemental Fund on High School Student Performance in Pilot Districts

September 2008

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

Carolina Institute
for Public Policy

The Impact of the Disadvantaged Student Supplemental Fund

on

High School Student Performance in Pilot Districts

GARY T. HENRY

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

CHARLES L. THOMPSON

EAST CAROLINA UNIVERSITY

C. KEVIN FORTNER

GEORGIA STATE UNIVERSITY

DANA K. RICKMAN

GEORGIA STATE UNIVERSITY

REBECCA ZULLI -LOWE

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Table of Contents

Summary of the Key Evaluation Findings

In its first two years of operation, the Disadvantaged Student Supplemental Fund (DSSF) distributed approximately $22.4M annually to 16 pilot districts selected on the basis of the level of educational disadvantage in each district. These districts received approximately $250 per pupil or $840 per academically disadvantaged pupil. Academically disadvantaged students are defined as those who did not achieve proficiency on either the 8th grade reading or 8th grade mathematics assessment. Based on a very rigorous research design and analysis, we conclude that the DSSF pilot program has had a positive impact on student achievement.

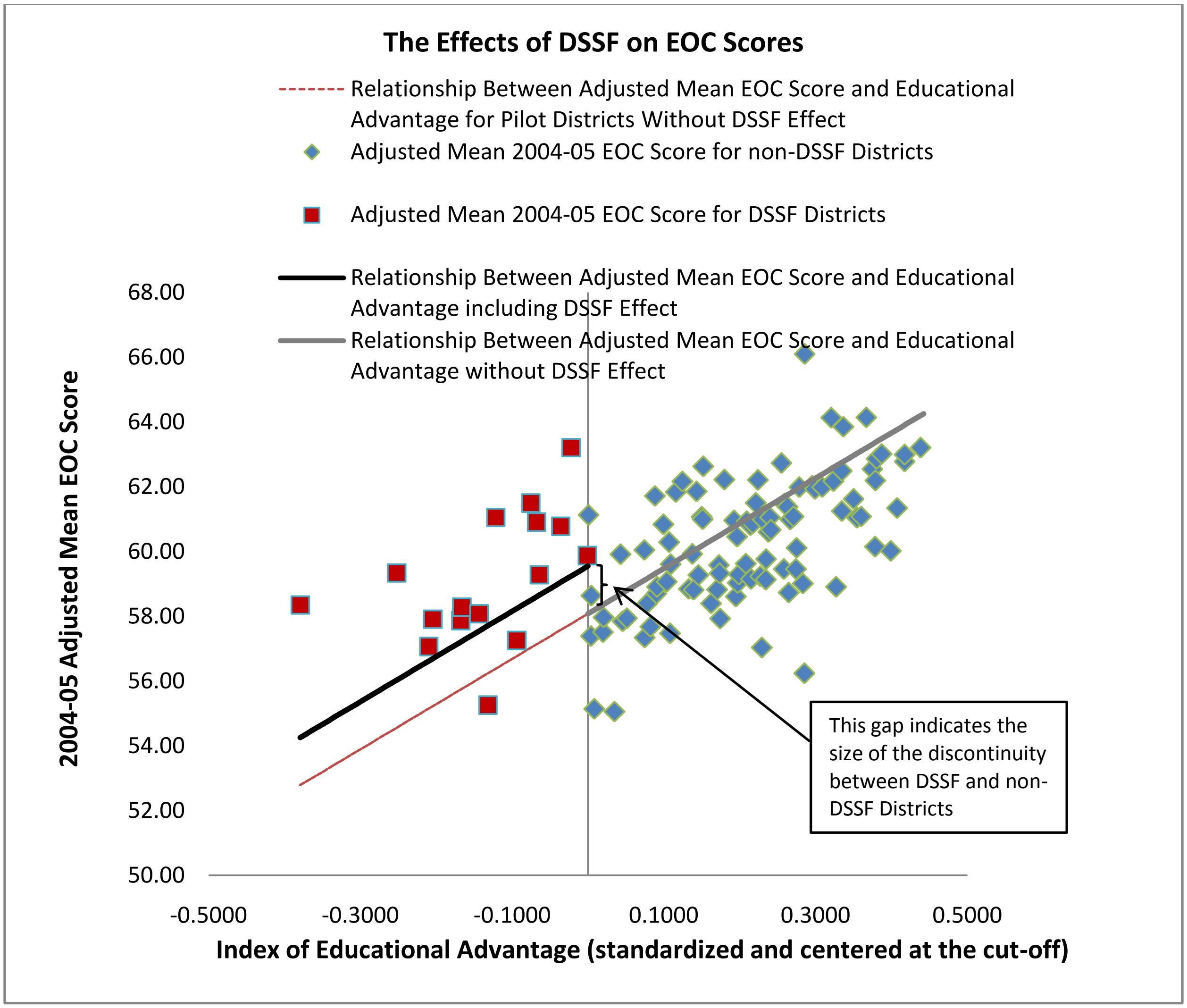
High school students in the 16 pilot districts scored significantly higher on their End-of-Course (EOC) exams than would have been expected in the absence of the program. Overall, high school students in the DSSF districts scored about 1.5 points higher on EOC exams than they would have scored without the program. The comparable gain for academically disadvantaged students was smaller – a little more than 0.5 points -- but still meaningful.

To get a sense of how meaningful these gains are, consider that with the DSSF program, the difference between average scores for all high school students in DSSF pilot districts and all high school students in other districts is about 3.7 points. Without DSSF, the expected difference in scores would be about 5.2 points. For just the educationally disadvantaged high school students, the difference in average EOC scores in DSSF versus other districts is about 0.4 points. Without the DSSF pilot, that difference would be more than twice as large, about 0.95 points.

But how can we conclude that it was the DSSF program that accounts for the 16 districts' better-than-expected performance? Districts were selected for the DSSF pilot strictly based on educational need, as measured by a single score combining four variables: the proportion of proficient students in the district, the percentage of teachers retained from year to year, the percentage of teachers with more than five years of teaching experience, and the percentage of the population living above the poverty line. To determine whether DSSF had an impact on student learning outcomes in pilot district high schools, we employed a very rigorous research design called regression discontinuity (RD). This design along with the proper analysis of the data, allowed us to control for other factors that might bias the estimates of the program's effectiveness and attribute gains to program impacts. The highly regarded Institute for Educational Sciences (IES) in the U.S. Department of Education recommends RD for situations in which random assignment to a program was not deemed possible; as was the case with the DSSF program. A full discussion of RD is offered in the appendix.

We graphically illustrate the research design and the impacts of the DSSF program for 2004-05 in Figure 1. On the horizontal axis, we plotted the educational need variable for districts in the state; and on the vertical axis, we plot their average EOC score adjusted for district differences. To the left of the vertical line the graph displays the DSSF districts and to the right the non-DSSF districts. The dotted line indicates the expected average EOC score, adjusted for district characteristics, that would have been expected in the absence of DSSF. The solid line indicates the expected average adjusted EOC score including the effect of the DSSF program. On the right side of the graph the dotted line and the solid line are the same. The difference between the two lines for the pilot districts indicates the size of the effect due to DSSF. The break in the solid line for the other districts and the DSSF districts indicates the program impact at the point that was chosen as the cutoff for receiving the supplementary funding. Clearly the discontinuity shows that the DSSF districts did outperform expectations, precisely at the cutoff point. After a very close inspection and several analyses, we find no other explanation for the sharp difference that occurs at precisely the cutoff point.

Figure 1:



This analysis shows that the DSSF pilot exerted its effects broadly rather than in a highly targeted manner – it improved performance for all students, not just for disadvantaged students. The DSSF was effective at reducing high school performance gaps between disadvantaged districts and other districts. The DSSF was also effective in reducing the gap between academically disadvantaged students in the pilot districts and similar students in other districts. It does not appear that DSSF was effective in reducing the gap between academically disadvantaged students and their more advantaged peers, however, the gains of the academically disadvantaged students did not come at the expense of proficient students. This has special significance in light of the Leandro school finance case. In that case, Judge Howard Manning observed pointedly and the Supreme Court has affirmed that the obligation to provide a sound basic education can be measured by the extent to which academically disadvantaged students' performance is raised to the level of competency but that should not come at the expense of the proficient students.

After finding that the DSSF pilot was effective in high schools, we attempted to determine the specific factors or strategies that led to the effects. From local interviews and state policy discussions several specific ways that DSSF might influence student outcomes have been mentioned, such as (1) more per total pupil spending; (2) more spending for regular instruction or for other functional categories; (3) more teacher compensation; (4) improved teacher quality; (5) decreased class size; (6) improved school leadership and the organizational conditions that such leadership can shape; (7) greater teacher motivation or effort in teaching the material on the standard course of study. Currently, we only have available data to assess the first five of these means for improving student performance. Despite the clear evidence of the DSSF pilot's positive impact, the first five explanations do not explain the effects. Existing data do not allow us to ascertain exactly how the program brought about the improvements in student performance. However, we are in the process of collecting data that should allow us to assess other potential mechanisms that may have produced the gains.

The balance of this report begins with a background discussion of the DSSF project, and then is followed by a short discussion of the project's data and methods. The final section details our findings to date and spells out the future steps required to identify the mechanisms through which the DSSF is working in the pilot districts.

Introduction

In 2004, Governor Mike Easley and the North Carolina State Board of Education established the Disadvantaged Student Supplemental Fund (DSSF) as a pilot program in 16 of the state's most educationally disadvantaged districts. The overarching goal of the program was to increase the learning and academic performance of students, especially academically disadvantaged students. During the 2004-05 school year, the program provided $22.4 million in additional funds to the pilot districts. The program allowed districts flexibility in using the funds to attract and retain qualified, competent teachers and to provide enhanced instructional opportunities to students at risk of school failure. The NC Department of Public Instruction (DPI) was required to provide assistance and monitor the program through the Local Education Agency Assistance Program (LEAAP). The pilot program continued with slightly increased funding for the 2005-06 school year. In 2006-07, the Governor recommended expanding the program statewide, and the General Assembly appropriated $49.5 million for this purpose. The original 16 pilot districts continued to receive the same amount of funding per pupil as they received in the first two years, but the remainder of the state's districts received approximately $88 per academically disadvantaged pupil.

In 2005, before the statewide expansion, the General Assembly mandated an evaluation of the effectiveness of the strategies funded by the DSSF program, as well as the effectiveness and efficiency of DPI's LEAAP teams. The team carrying out the evaluation is based primarily at the University of North Carolina at Chapel Hill, with participation by researchers at East Carolina University and Georgia State University. This is the third in a series of reports from the research team, but the first to report findings on the impact of the DSSF on student performance. The first report compared the quality of teachers available to disadvantaged and other students in pilot districts with the quality of teachers available to students statewide, traced patterns in the performance of students progressing through NC schools, described how pilot districts spent their funds and implemented the program, and explored the strengths and limitations of assistance provided by DPI's LEAAP teams. The second report provided an update on teacher quality in pilot districts during the second year of the program and assessed the use of DSSF to supplement teachers' salaries.

As implied in the summary of key findings above, the current report was designed to address three questions:

(1) Did the DSSF program improve the performance of high school students in pilot districts during the first two years of the program?
(2) More specifically, did the DSSF program improve the performance of academically disadvantaged high school students in these districts – students who entered the ninth grade with reading and mathematics skills below the "proficient" level on the 8th Grade EOG tests?
(3) If the program did improve high school student performance in participating districts, what specific mechanisms or strategies produced the gains?

Next we describe the methods we used to address the questions. Then we present our findings in detail.

Data and Methods

The data required for our analyses -- an extensive array of school, classroom, teacher, and individual student characteristics for the 2004-05 and 2005-06 school years – were provided by the NC DPI. Although our analyses focused on the impacts in the DSSF pilot districts, these results include all of high school students, classrooms, and teachers from all of the 330 regular high schools that were in operation in both study years. The variables we selected for use in our analyses have all been shown to affect student outcomes in other contexts. For details on the regression discontinuity design, data, variables, and the hierarchical linear models used for the analysis, see the Appendix to this report.

In the findings sections of the report we describe our estimates of the "net effect" of the DSSF program on high school student achievement. This model provides an estimate of the effect of DSSF participation and funding on student achievement, leaving out variables that might have been affected by the DSSF pilot such as teacher quality, class size and per pupil expenditures. Because DSSF provides money to support teacher bonuses, which might result in higher quality teachers, technology, and other types of expenditures, positive effects of DSSF could be hidden if we included variables whose values changed due to receipt of DSSF funds. We also used existing administrative data to try to explain *how* DSSF produced the net effects. In these subsequent models, we do include variables such as teacher quality, class size and per pupil expenditures in analyses to attempt to reveal the mechanisms of *how* DSSF worked. In both types of analyses, we employed Hierarchical Linear Modeling (HLM). HLM correctly adjusts for "nested" data. In our case, we have students nested within classrooms, which in turn nest within schools. For more details, including the equations for both types of models using HLM, see the Appendix.

Findings

In this section, we first provide details on our finding that the DSSF pilot did improve student performance in participating districts. We then turn to the question of *how* the program brought about this effect.

Impact of the DSSF Pilot on High School Student Performance: Discussion

We used two different variables to test the effect of DSSF on high school student performance in pilot districts. The first variable was an indicator of whether a district was a DSSF pilot district or not. This indicator enabled us to separate DSSF from non-DSSF pilot districts and compare students' average EOC scores between the two sets of districts. The second variable was the level of DSSF expenditures on a per-pupil basis. Including the second variable in an analysis could add to our confidence in the findings, if they are both positive since it would indicate that the amount of DSSF funds spent in the high schools relates to higher achievement. In both cases, we included the assignment variable used to select the pilot districts and other controls, which greatly increases our confidence that the estimate of the program effect can be attributed to DSSF.

We focus on two groups of students in this report. The first group consists of all of the students attending one of the state's 330 regular public high schools. The second group is the academically disadvantaged students in the same high schools. Judge Manning and the North Carolina Supreme Court have agreed that the percentage of students who achieve proficiency on state tests[1] is one measure of whether the state is meeting its obligation to provide a sound basic education. One purpose of the DSSF pilot was to increase achievement among the students who have been below proficiency in an earlier grade. Therefore, in our analysis we define academically disadvantaged students as those not meeting proficiency in either reading or mathematics at the end of the 8th grade.

In addition, we included a number of school, classroom, and individual student "control" variables along with the DSSF variables in our initial analyses. The control variables were measures of factors that were not directly affected by the program but have been shown to affect student achievement. These models yielded an estimate of the "net effects" of DSSF – that is, the effects of DSSF after separating out the effects of the school, classroom, and individual level variables that are beyond the scope of DSSF.

The left-hand column in Table 1a below displays the results from the "net effects" model based on the simple distinction between DSSF pilot districts and other districts.

Table 1a: Estimates of DSSF Program Effects for All Students and Academically Disadvantaged Students

Net Effects	**EOC Score - All Students**		**EOC Score - Academically Disadvantaged Students**	
	Coefficient	Std. Err	Coefficient	Std. Err
DSSF Indicator	**0.1646**	**(0.0360)**	**0.0866**	**(0.0380)**
School Level				
Percent Asian	-0.0018	(0.0025)	**-0.0064**	**(0.0027)**
Percent Black	0.0006	(0.0003)	**0.0008**	**(0.0004)**
Percent Hispanic	0.0019	(0.0017)	0.0020	(0.0017)
Percent Multiracial	**0.0086**	**(0.0042)**	-0.0038	(0.0053)
Percent American Indian	-0.0011	(0.0008)	-0.0009	(0.0007)
Percent Free lunch	-0.0012	(0.0006)	-0.0009	(0.0007)
Percent Reduced lunch	0.0008	(0.0014)	0.0014	(0.0017)
School size	0.0000	(0.0000)	-0.0000	(0.0000)
Indicator for 2005-06	**-0.0377**	**(0.0104)**	-0.0138	(0.0117)
Educational Need	**0.4711**	**(0.1237)**	0.1509	(0.1483)
Educational Need Squared	**0.5243**	**(0.2240)**	-0.0342	(0.4341)
Educational Need Cubed	**-2.6691**	**(0.7319)**	0.1280	(1.1572)
Classroom Level				
Advanced curriculum	**0.1407**	**(0.0094)**	**0.2114**	**(0.0243)**
Remedial curriculum	0.0259	(0.0362)	0.0324	(0.0425)
Classroom Ability Dispersion	**0.0333**	**(0.0138)**	**-0.0631**	**(0.0253)**

[1] In the ABCs assessment system, student scores on EOG and EOC tests fall into four levels of achievement. The knowledge and skills demonstrated by students scoring that are not adequate to enable them to succeed in the next grade in school. Not proficient is divided into two levels: Level 1 (insufficient mastery) and Level 2 (inconsistent mastery), and proficient students receive either a Level 3 (consistent mastery) or Level 4 (superior mastery) score.

Table 1b: Estimates of the Impact of the DSSF Pilot Program: Individual Student Level Characteristics

Net Effects - All Students	**DSSF All Students**		**DSSF Academically Disadvantaged Students**	
	Coefficient	Std. Err	Coefficient	Std. Err
Individual Level				
8th Grade Reading Score	**0.2481**	**(0.0020)**	**0.2368**	**(0.0051)**
8th Grade Mathematics Score	**0.4527**	**(0.0028)**	**0.3092**	**(0.0065)**
Absences	**-0.0087**	**(0.0001)**	**-0.0067**	**(0.0003)**
Classroom peer ability	**0.1742**	**(0.0075)**	**0.2010**	**(0.0109)**
Male (=1)	**0.0254**	**(0.0023)**	-0.0056	(0.0061)
Asian (=1)	**0.0447**	**(0.0078)**	0.0549	(0.0331)
Black (=1)	**-0.0987**	**(0.0034)**	**-0.1109**	**(0.0080)**
Hispanic (=1)	-0.0064	(0.0064)	0.0061	(0.0184)
Multiracial (=1)	**-0.0179**	**(0.0071)**	-0.0002	(0.0242)
American Indian (=1)	**-0.0634**	**(0.0101)**	**-0.0860**	**(0.0208)**
Eligible for free lunch	**0.0158**	**(0.0027)**	0.0010	(0.0072)
Eligible for reduced price lunch	**0.0185**	**(0.0040)**	**0.0230**	**(0.0117)**
Eligibility free/reduced missing	**0.0325**	**(0.0164)**	-0.0152	(0.0395)
Gifted	**0.1079**	**(0.0045)**	0.0326	(0.1028)
High incidence	**-0.0452**	**(0.0053)**	**-0.0776**	**(0.0092)**
Cognitive	**-0.1325**	**(0.0171)**	**-0.2059**	**(0.0181)**
Behavioral	-0.0078	(0.0195)	-0.0209	(0.0239)
Sensory	-0.0215	(0.0236)	-0.0449	(0.0491)
Physical	**-0.0840**	**(0.0116)**	**-0.1067**	**(0.0160)**
Severe	**0.0396**	**(0.0117)**	0.0031	(0.0243)
Previous LEP	-0.0190	(0.0144)	-0.0189	(0.0237)
Current LEP	**0.0424**	**(0.0183)**	**0.1391**	**(0.0539)**
Parent Ed: <HS	**-0.0235**	**(0.0046)**	**-0.0545**	**(0.0107)**
Parent Ed: = HS	**-0.0221**	**(0.0024)**	**-0.0483**	**(0.0069)**
Parent Ed: = College	**0.0146**	**(0.0027)**	-0.0104	(0.0081)
Parent Ed: Missing	**-0.0915**	**(0.0177)**	**-0.1203**	**(0.0311)**
Underage	**0.1059**	**(0.0075)**	**0.1318**	**(0.0294)**
Overage	**-0.0788**	**(0.0028)**	**-0.0860**	**(0.0061)**
Grade 9	0.0048	(0.0054)	**0.0483**	**(0.0117)**
Grade 10	0.0086	(0.0073)	**0.1011**	**(0.0131)**
Grade 11	**0.0846**	**(0.0094)**	**0.1687**	**(0.0173)**
Grade12	**0.0386**	**(0.0101)**	**-0.0272**	**(0.0138)**
English 1	**-0.3107**	**(0.0117)**	**-0.3050**	**(0.0181)**
Algebra 1	**0.1827**	**(0.0124)**	**0.0462**	**(0.0154)**
Algebra 2	**-0.0568**	**(0.0109)**	**-0.1111**	**(0.0142)**
Geometry	**-0.5694**	**(0.0161)**	**-0.6461**	**(0.0302)**
Physical Science	**-0.9870**	**(0.0282)**	**-1.0897**	**(0.1357)**
Biology	**-0.2567**	**(0.0101)**	**-0.3655**	**(0.0160)**
Intercept	**-0.0976**	**(0.0215)**	-0.0333	(0.0308)

Table 1a on page six indicates that after separating out the effects of all of the control variables, schools in DSSF districts produced average EOC scores that were approximately 0.16 of a standard deviation higher than would have been expected. As noted previously, this translates into an overall difference of about 1.5 points in average EOC scores. The DSSF high schools produced better average EOC scores than would have been expected if the supplemental funds were not made available to the pilot districts.

Not shown in a table, the other way that we measure the impact of DSSF – the per pupil DSSF Expenditure at each school – enables us to estimate how much better DSSF students did for each $100 of per-pupil funding provided by DSSF. The results show that $100 in DSSF funds produced an increase of about 0.05 of a standard deviation in average scores. The average DSSF expenditure for all schools in the DSSF pilot districts was about $250 per student. The effect of the DSSF $250 was about 0.1325 of a standard deviation, which approximates the 0.16 SD effect from the DSSF indicator variable noted above. This helps to corroborate the first analysis.

Finally, in Table 1a, we show that the program effect for disadvantaged students is 0.087 standard deviations or nearly 0.55 points on the average EOC exam. The DSSF program effect for the academically disadvantaged student is about one-third of the level of the average student. This indicates that the likely effect of adding more resources to districts with high educational need is an overall increase in achievement which also raises the achievement of academically disadvantaged students, but to a lesser extent. These funds appear to improve achievement across the board by raising the performance of both the lower performing students and the higher performing students. The DSSF program effect does not appear to reduce the gap in performance, between disadvantaged students and all other students. In fact, the DSSF program appears to increase the gap.

The Effects of Control Variables: Individual Student Level

Since achievement on the EOC tests also depends on student, classroom, and school characteristics, we have included variables to account for these differences. These variables are displayed in Tables 1a and 1b to separate their contributions to student performance from the contribution of the DSSF pilot. This allows a more accurate estimate of the effects of the DSSF. We mention these results briefly to provide a few insights about the effects of these variables which are useful in their own right.

At the individual student level, a variety of student characteristics exerted a statistically significant effect on a student's average EOC scores. Among individual characteristics, how a student performed on his or her 8th grade EOG assessments in reading and mathematics are the best predictors of that student's performance on EOC exams in later grades. Scoring one standard deviation higher on 8th grade EOG reading raises the average outcome on EOC exams by about 0.25 standard deviations, holding other variables constant. An increase of one standard deviation on a student's 8th grade EOG math score is expected to increase the student's EOC test score by about 0.45 standard deviations for students similar in all other respects. The number of days a student is marked as absent from school has a negative relationship with EOC test scores. An additional 10 absences lowers student performance by about 0.08 standard deviations, setting aside the effects of other variables. The average student was marked absent from school approximately 8 days during the school year.

The average ability of a student's peers in their high school classes (measured in 8th grade EOG reading and mathematics scores) has a positive relationship with students' EOC test scores. Students in a classroom with other students whose scores average one standard deviation above the average score perform about 0.17 standard deviations higher than similar students with peers whose EOC scores are average. Being in a class with more able peers improves each high school student's scores. However, the classes tend to have scores that are closer to average, which makes the average increment closer to 0.03 standard deviations or about 3/10th point on average.

On average, male students score slightly higher than comparable female students, by about 0.025 standard deviations. Holding constant other factors, Asian and multi-racial students perform slightly better (0.04 and 0.02 standard deviations respectively), black and American Indian students perform slightly worse (-0.09 and -0.06 standard deviations respectively), while Hispanic students perform no different statistically when compared with White students.

EOC testing outcomes also vary by income level. Perhaps surprisingly, compared to students not eligible for federal lunch subsidies, students eligible for either free lunch, reduced price lunch, or students whose free or reduced lunch eligibility data was missing all performed slightly better (about 0.02, 0.02, and 0.03 standard deviations respectively). In most studies, poverty as indexed by eligibility for federal lunch subsidies tends to reduce low-income students' test scores. Our counterintuitive result probably reflects two facts. First, the low reading and math skills that many low-income students bring into high school have already been separated out in our analyses. Second, many low-income students drop out during high school, and those who remain may be more highly motivated. The combination of these two factors probably explains our unusual result.

The breadth of the data available for this analysis allowed us to break out students with various exceptionalities in our analyses. Students coded as academically or intellectually gifted score on average about 0.11 standard deviations higher than comparable regular instruction students (those coded as neither gifted nor disabled) on EOC exams. Within other exceptionality categories, performance on EOC exams is generally lower compared to otherwise similar regular instruction students. Students with cognitive disabilities, high incidence disabilities, and those with physical disabilities scored lower on average than comparable regular instruction students (about 0.13, 0.04, and 0.08 standard deviations respectively). Students coded as having behavioral or sensory disability performed as well as students with no coded disabilities. Finally, students coded as having a severe disability actually performed slightly better on average than comparable regular instruction students. This finding is likely attributable to selection bias where only higher performing students with this type of disability were actually tested, or administrative errors in coding given a very small number of students in this category.

Students coded in the data as receiving Limited English Proficiency (LEP) services performed about the same as comparable students never served by LEP programs. Students served by LEP in previous years, however, score on average about 0.04 standard deviations higher than comparable students on EOC tests. This finding may seem contrary to expectations, but keep in mind that a student's 8th grade performance is included in the model. As these students become more proficient in the English language, they might improve at a faster rate than comparable students, which would explain this difference in achievement.

The level of education of a student's parents was also a statistically significant predictor of student achievement on EOC tests. Otherwise comparable students whose parents attended some college performed slightly better than students whose parents did not finish high school or were high school graduates (about 0.02 standard deviations on average for both groups). Students whose parents were college graduates performed, on average, slightly better than those whose parents attended some college (0.015 standard deviations). Students whose parental education level was missing in the data scored about 0.09 standard deviations lower than students whose parents attended some college.

Using a student's date of birth, students were coded into three categories: under age, over age, and of age for grade level. The coefficients for the under age and over age variables indicate the difference in EOC test achievement between similar students who differ on only the under age or over age variable. Under age students, those whose birthdates/age would place them in a lower grade level (they may have started school early or skipped a grade), scored about 0.11 standard deviations higher compared to similar students of the appropriate age for their grade level. Over age students, those whose birthdates would place them in a higher grade level given their age (they may have started school late or been retained previously), scored about 0.08 standard deviations lower compared to similar students of the appropriate age for their grade level.

Controls for grade level indicate that students in grades 9, 10, and 11 score about the same on EOC exams controlling for differences in other characteristics. Students in grade 12 score about 0.08 standard deviations higher on average compared to similar students in other grades.

The final group of controls adjusts the results to account for differences between EOC exams. The reference group is Algebra I, and coefficients indicate the average difference in achievement between students enrolled in the specified course and comparable Algebra I students. For example, students in English I scored about 0.03 standard deviations higher than comparable students in Algebra I. The purpose of including this set of controls is to ensure that our results – which are expressed in terms of average EOC scores, overall – are not biased by differences in the difficulty levels of different exams, the particular populations of students taking different exams, or other factors associated with the individual exams.

The Effects of Control Variables: Classroom Level

Enrollment in a class with an advanced curriculum – honors or AP -- has a positive impact on EOC test scores, holding prior achievement and other control variables constant. In other words, the challenging nature of the curriculum is likely to have produced this effect, since the prior ability level of the student is controlled. Students in classes with advanced curricula score about 0.14 standard deviations higher than similar students attending classes with regular curriculum. In addition, having a more diverse group of abilities within a classroom is associated with higher test scores, holding other characteristics constant. A student enrolled in a class with students who are one standard deviation more diverse scores on average about 0.03 standard deviations higher on his or her EOC tests than a comparable student in a classroom where every student is of the same ability (standard deviation of zero).

The Effects of Control Variables: School Level

In the foregoing discussion, we have generally limited ourselves to commenting on control variables which proved statistically significant. But particularly in the case of certain school level variables, it is worth pointing out that some variables found significant in other studies were *not* significant in the present analysis. Once the effects of all of the other variables already discussed are taken into account, the ethnic composition of a high school's student population did not generally exert a significant effect on average EOC scores, nor did the percentage of students receiving free or reduced price lunch. Controlling for all other variables in the model, the size of the high school did not affect average EOC scores in the school. Note also that controlling for other variables, in academic year 2005-06, average EOC scores were a little lower than average EOC scores in the 2004-05 school year. All of these findings apply statewide, not solely in DSSF pilot districts.

Finally, the "educational need index" variable is the index used to select the school districts included in the DSSF pilot. The four variables in the index were the proportion of proficient students in the district, the percentage of teachers retained from year to year, the percentage of teachers with more than five years of teaching experience, and the percentage of the population living above the poverty line. By including the index in our model, we separate out the effects of these variables from the effects of the DSSF program itself upon high school student performance in the selected districts. For technical reasons that relate to the nature of the relationship between the index and EOC scores, we have included squared and cubed terms for the index. Including these terms provides stronger evidence that the estimated effects of DSSF are accurate or, in more technical terms, unbiased.

How Did DSSF Improve Student Performance?

Many different variables might explain *how* the DSSF pilot improved high school students' performance in these districts. From local interviews and state policy discussions several specific ways that DSSF might influence student outcomes have been mentioned, such as (1) increasing total per pupil spending; (2) increasing spending for regular instruction or for other functional categories; (3) increasing teacher compensation; (4) improving teacher quality; (5) decreasing class size; (6) improving school leadership and the organizational conditions that such leadership can shape; (7) increasing teacher motivation or effort in teaching the material on the standard course of study. Data available from the NC DPI permitted us to estimate the effects of the first five sets of variables, but not the last two. That is, data on school leadership practices and related school conditions or teacher motivation and effort are not currently collected in a way that allows accurate estimations of their effects on student achievement.

Our analyses involving the first five sets of variables indicate that some of these variables do exert significant effects on student performance, but none of the variables explains the DSSF pilot program effect. If any of these variables did explain the DSSF effect, then when they were introduced into our calculations, they would reduce the direct effect of the DSSF variable. This would suggest that per pupil expenditures, teacher compensation, teacher quality, or class size variables are mechanisms through which the program exerted its effect. But when we added each of them in turn to the net effects models, the coefficients that estimate the size of the DSSF effect were left virtually unchanged indicating that while these variables in some cases do influence student achievement, they do not explain the positive DSSF effect. We present our findings on these variables in the next section, and then we will comment about the alternative explanations we are now exploring.[2]

[2] Although tested, results for the impact of total per pupil expenditures are not presented here. The findings were consistent with all other tested explanations of the DSSF net effect.

Now we consider each set of variables in turn. It is important to remind the reader that the effects we report for each variable are independent or "net" of the effects of all of the other variables in our models, including the control variables listed in Tables 1a and 1b.

Expenditures by Function. As Table 2 indicates, the level of *regular classroom instruction* did not exert a significant effect on the average EOC scores of students overall, but higher spending on regular classroom instruction did correlate with higher average EOC scores for disadvantaged students. An additional $1000 spent on regular instruction translated into about 0.3 additional points on educationally disadvantaged students' average EOC scores – a modest but significant effect. In an earlier study, we did find that expenditures for regular instruction had a significant positive effect on EOC performance for all students, but our present model controls for the assignment of teachers to particular classes of students (Henry, Thompson, and others, 2008). The difference between the two studies is likely to be explained by the fact that higher performing students get the most experienced and most highly qualified teachers which is already taken into account in the estimates for this study.

It is actually not surprising that additional regular classroom expenditures have an effect on educationally disadvantaged students but not on students in general, even after controlling for the teacher-student assignment to classes. As we expressed the matter in our report on the High School Resource Allocation study, two broad types of resources shape how much students learn: the resources that students bring to school, and the resources that the school brings to students. By definition, educationally disadvantaged students bring lower reading and mathematics skills with them into high school. So it would be reasonable to expect that high schools would need to provide them with more instructional resources than other students would need in order to reach the threshold required to support proficient performance in high school courses. Our results show that when high schools do provide the additional resources for regular classroom instruction, those additional resources do pay off.

Table 2: Explaining the Effects of DSSF on High School Student Achievement: Class Size; Per pupil Expenditures; teacher Quality; and Teacher Pay

	All students Coefficient	**EOC IMPACT**	**Disadvantaged Students Coefficient**	**EOC IMPACT**
Net Effects Model- DSSF Indicator Coefficient -No Mediators	**0.1646****	1.4751	**0.0866***	0.5545
DSSF Indicator Coefficient for Mediator Class Size	**0.1651****	1.4797	**0.0867***	0.5555
Class Size	**-0.0015****	-0.0142	-0.0002	
DSSF Indicator Coefficient for Mediator Expenditures (in hundreds)	**0.1570****	1.4069	**0.1052****	0.6737
Regular Classroom Instruction	0.0019		**0.0045****	0.0290
Special Instruction	0.0051		0.0039	
Professional Development	**-0.0723****	-0.6482	-0.0309	
Supplemental Instruction	**-0.0156***	-0.1401	**-0.0221***	-0.1415
Student Services	-0.0081		-0.0077	
Technology	0.0070		0.0003	
School Leadership	-0.0078		0.0013	
LEA Instruction Expenditures	-0.0000		-0.0000	

	All Students Coefficient	EOC IMPACT	Disadvantaged Students Coefficient	EOC IMPACT
DSSF Indicator Coefficient for Mediator Teacher Pay (in thousands)	**0.1694****	1.5181	**0.0906***	0.5798
Teacher Salary	**0.0008****	0.0078	0.0004	
Teacher Bonus	**0.0077****	0.0690	0.0060	
DSSF Indicator Coefficient for Mediator Teacher Characteristics	**0.1583****	1.4186	**0.0812***	0.5201
National Board Certification	**0.0423****	0.3794	**0.0440****	0.2817
Advanced Degree	-0.0067		-0.0143	
Most & Highly Comp. Barron's	0.0124		0.0137	
Teacher Test Average	**0.0121****	0.1087	0.0044	
First Year Teacher	**-0.0899****	-0.8059	**-0.0677****	-0.4334
1-2 years experience	**-0.0179***	-0.1606	-0.0239	
3-4 years experience	0.0016		0.0074	
Comparison group is 5-9 year of experience	---------------		------------------	
10-14 years experience	-0.0122		-0.0212	
15-19 years experience	-0.0048		-0.0113	
20-24 years experience	-0.0175		**-0.0440****	-0.2821
25 plus years experience	**-0.0163***	-0.1468	-0.0120	
Infield Teacher	**0.0434****	0.3894	**0.0448****	0.2871

More dollars spent on supplemental instruction – instruction given after school, on weekends, and in the summer – do not improve the performance either of all students or of educationally disadvantaged students on the EOC exams. In fact, more spending on supplemental instruction is significantly correlated with lower average EOC scores. It seems illogical to infer that spending more on supplemental services lowers achievement. It is much more likely that schools with substantial numbers of low-scoring students try to improve their performance via supplementary instruction, but the supplementary instruction is not enough to raise the test scores a significant amount to offset the larger number of students requiring the services. These findings do not mean that supplemental instruction cannot or does not pay off in some settings. Some schools or supplemental instruction providers may have figured out how to carry out supplemental instruction in an effective way. But on average across NC high schools, how to provide effective supplemental instruction is an unsolved problem.

Similarly, more spending on student services fails to increase EOC scores for either students in general or educationally disadvantaged students in particular. One might not necessarily expect student services such as counseling, psychological screening, testing for various disabilities, and the like to yield a discernible effect on average EOC scores in a school. In some cases, they target a very small number of students. In others, they are only indirectly connected with academic learning. Such services may serve other important functions, such as keeping some students in school, putting them on appropriate paths to postsecondary education, or addressing social needs outside of school. We simply note that more spending on student services does not produce higher EOC scores.

For the general student population, more spending on professional development is correlated with lower EOC scores, but the association probably means that the allocation of more dollars is not yielding returns in EOC score improvements. More spending on professional development does not pay off for disadvantaged students, either. Finding that professional development does not improve student achievement is fairly commonplace in recent research, but these results do not mean that PD cannot improve student performance. Currently, on average, high schools have not been spending professional development dollars in a way that increases achievement.

More spending on the other functions displayed in the table also does not contribute significantly to average EOC scores for either students in general or for educationally disadvantaged students in particular after the other controls are taken into account. Finally, note that even taken as a set, the expenditure variables have virtually no impact on the estimate of the effect of the DSSF pilot program.

Teacher Salaries and Bonuses. Higher spending on either teachers' salaries or the supplements and bonuses they earn – including but not limited to the ABCs bonuses – have only a negligible impact on overall EOC scores in a high school, and their effects on educationally disadvantaged students' EOC scores is essentially zero. These variables did not explain the positive effect of the DSSF pilot.

Teacher Quality. Several of the teacher quality variables in our model had a significant impact on students' average EOC scores. The strongest effects were those for National Board Certified Teachers and teachers who are licensed to teach the specific EOC-tested subjects they were teaching. Both all students and educationally disadvantaged students taught by an NBCT perform substantially better than do similar students taught by a teacher who is similarly qualified in all other measured effects but who is not National Board certified. The coefficient for National Board certification translates into about 0.38 points in a student's average EOC scores.

The effect of having a teacher who is licensed to teach the EOC-tested subjects s/he is assigned to teach is similar to that for National Board Certification. The coefficient for in-field teaching also translates into a difference of about 0.39 points on a student's EOC exams, and this applies to both all students and to educationally disadvantaged students.

The higher a teacher has scored on the various academic and professional tests in his or her licensure file, the better overall student performance s/he produces, but this variable does not have a significant impact on EOC scores for educationally disadvantaged students. Examples of the scores averaged to create this variable are SAT or ACT scores, score on the National Teacher Examination, and all types of PRAXIS scores. (Scores were standardized to make them comparable.) A one standard deviation difference in average test scores translates into about a 0.11 point better EOC scores for that teacher's students, overall. Once the effects of all other variables are removed, teachers with advanced degrees produce no better EOC scores than do other teachers.

Our final subset of teacher quality variables relate to teaching experience. Here we have chosen to compare sets of teachers with different levels of experience. A good deal of research shows that teachers have generally reached the height of their powers by the time they have five to nine years of experience, and our results are consistent with that finding. Compared to teachers in this category, teachers with less than 3 years of experience are less effective. Students of teachers in their first year of teaching score about 0.8 points lower on EOC exams than students of teachers with 5 to 9 years of experience. Things improve for teachers in their second and third years. Their students score only about 0.16 points lower. Between 3 years of experience and 24 years of experience, there is little difference – their students score about the same on EOCs. But after 25 years, a teachers' impact on EOC performance tends to decline. Our interviews for the High

School Resource Allocation study suggest that the decline may reflect more experienced teachers' negative response to the pressures of accountability systems. Younger teachers tend to regard accountability pressures as a simple fact of life. Other possibilities are that more experienced teachers' subject matter knowledge becomes outdated, that they fall out of step with their adolescent students, or that their levels of energy and commitment begin to decline as they approach retirement.

Taken together, the effects of the statistically significant teacher quality variables is quite substantial. Consider a student who is taught by a National Board Certified Teacher, teaching in-field, with a record of test scores one standard deviation above the mean, with 5 to 9 years of experience. Compare that student's average EOC performance with a similar student taught by a teacher with none of these advantages and in his or her first year of teaching. On the basis of teacher quality variables alone, the first student would average about 1.67 points better on the EOCs than the second. To put this into perspective, the difference between average EOC score in the top 25% of the state's high schools and average EOC scores in the bottom 25% is only about 5-7 points.

One major alternative to these potential explanations for the impact of the DSSF pilot is the possibility that district leaders, school leaders, and teachers simply knew that they had been singled out for extra funding and extra attention and therefore worked harder and smarter to produce more student learning and better student performance. This interpretation is consistent with our finding that the effects of DSSF are global rather than specific to disadvantaged students. But this interpretation is not as simple as it first appears. If educators in these districts worked harder and smarter, in what ways did they do so? What did this stronger motivation and smarter practice look like?

Our research on a related project – the High School Resource Allocation study commissioned by Governor Easley – suggests that principals' leadership practices and the organizational conditions they cultivate may be the keys to improved performance. Drawing upon our own and others' research on leadership and associated factors, we constructed a questionnaire designed to measure several key leadership-related variables. These include:

1. teachers' commitment to the organization and its goals;
2. trust between teachers and principals as well as among teachers;
3. professional community;
4. the extent to which teachers actually teach the content of the NC Standard Course of Study;
5. the extent to which teachers collaborate with the principal to monitor student progress, and;
6. the degree to which principals actively work to promote these conditions.

As a result of the findings in our studies of high schools, we added collection of these data as we began our focus on middle schools. Data collection in middle schools in DSSF pilot districts and in a matched set of similar middle schools in neighboring districts is almost complete. We expect to process and analyze the data during the summer and report our analyses of them in the fall.

Although the expenditure, teacher compensation, and teacher quality variables do not explain how the DSSF pilot improved the performance of high school students in those districts, our findings about the effects of each do illuminate the dynamics underlying high school student performance statewide.

Conclusion

The analyses in this report leave little doubt that the DSSF pilot program produced positive effects on high school achievement overall and specially for academically disadvantaged students. The program effect is estimated to be 1.5 points on EOC tests and 0.5 on the EOC tests for students who reached high school without proficiency on at least one of their 8th grade tests. The gains that we observed in the pilot districts did reduce the gap of these districts with significant educational needs with other districts, but in these districts the gains of the academically disadvantaged students did not come at the expense of higher performing students. In fact, the program's effects on the higher performing students were nearly three times as large as the program's effects on the academically disadvantaged students.

The strength of these claims for the program's effects should not be underestimated. During the last several years, the bar for "scientifically based research" to estimate educational program effects has been significantly raised. This evaluation used one of the three types of design – regression discontinuity – that are frequently singled out for producing unbiased estimates of program effects. In addition to using a very strong design, we have incorporated state-of-the-art hierarchical linear modeling. Finally, we ran the analysis in several different ways to ensure that the findings were not due to the educational needs of the districts in the pilot or other factors. Each analysis produced similar estimates of the program effects and our confidence in the program effects were enhanced. In short, we found no evidence that diminished our confidence in the findings. We also did not find an alternate explanation for the program effect that is observed at precisely the cut-off point that was used to choose the pilot districts.

The fact that the program effect occurs precisely at the cut-off point is a strong indicator that these districts did not simply rebound from previously lower than normal EOC scores or that districts with the greatest educational need all do better. In spite of the fact that the findings strongly point to significant effects of DSSF on high school achievement, currently available data do not explain how the effect has occurred. In future studies, we will be able to delve deeper into the explanations. In our evaluation of DSSF effects on middle school achievement, we will include an analysis of the effects of leadership and content coverage in mathematics and reading as potential explanations of the effects. Once we have more years of data, we can take a look at the effects of specific changes in funding and allocation of other resources in an attempt to provide explanations for the effects of the DSSF program. For now, we will have to be satisfied that the DSSF funding targeted to North Carolina school districts with high levels of educational need has a positive affect on high school achievement.

Appendix 1: Study Data and Methods

DATA

The primary data for the study combined student, classroom, teacher, and school variables provided by the North Carolina Department of Public Instruction (DPI) for the 2004-05 and 2005-06 school years.

2004-05 and 2005-06 Student Data

Student data was provided by the North Carolina DPI from administrative End-of-Course test records and school roster reports for 2004-05 and 2005-06. This data included information on the student's previous end of grade test history, demographic information, exceptionality codes, and any current or previous classification as a limited English proficiency student (LEP). The student level data on date of birth was used to compute an indicator for students that are over age or underage given their grade level and the cutoff date for enrollment into public school. Lacking specific data on a student's grade retention in the past, this overage measure provides a proxy for student retention in previous years. A measure of peer ability is calculated for each student using the 8th grade end of grade tests scores for other students in each student's EOC class.

From previous test score data, calculations were also made to determine if a student was considered 'academically disadvantaged' based on their 8th grade End-of-Grade (EOG) reading or mathematics score.[3] Academically disadvantaged high school students are students who scored not-proficient (either level 1 or level 2) on either their 8th grade reading or mathematics End-of-Grade (EOG) assessment. The knowledge and skills demonstrated by students scoring at these levels are not adequate to enable them to succeed in the next grade in school on either their reading or math tests. They have not reached grade level "proficiency." For example, students who score at Level 1 or Level 2 on the 8th grade assessments have not attained sufficient mastery of eighth grade reading or mathematics skills to enable them to succeed in high school. By contrast, students scoring at Levels 3 or 4 have developed the skills necessary to succeed in the next grade. They have reached "proficiency."

Table A.1 provides a detailed description of the data for a) the full high school student population included in the sample and, b) the population of disadvantaged students. The number of cases listed is weighted based on the number of semester observations included in the data during the two year period. Each observation represents a student in a semester enrolled in an EOC tested subject. For example, a student taking Algebra I and English I for the 2004-05 school year and enrolls in Biology I for the 2005-06 school year is listed in the data six times (two semesters per class per year).

It is worth noting the differences between the full population of students and the population of academically disadvantaged students. All scores listed within tables and included in the analysis are within-year standardized scores based on all valid test scores included in the provided data. With standardization, the average score becomes zero and a score of one indicates a student who

[3] In the ABCs assessment system, student scores on EOG and EOC tests fall into four levels of achievement. The knowledge and skills demonstrated by students scoring that are not adequate to enable them to succeed in the next grade in school. Not proficient is divided into two levels: Level 1 (insufficient mastery) and Level 2 (inconsistent mastery), and proficient students receive either a Level 3 (consistent mastery) or Level 4 (superior mastery) score.

scored one standard deviation above the mean (or average) score. In the full sample, grade 8 reading and math scores are slightly above average indicating that students with higher than average performance in 8th grade are likely to attempt more EOC tested courses. Disadvantaged students scored on average more than one standard deviation lower compared to the full sample. This is as expected as the indicator for academic disadvantage is based on 8th grade EOG scores.

The academically disadvantaged subpopulation has a higher proportion of Black students, students eligible for free or reduced priced lunch, and students whose parents either did not complete high school or only have a high school diploma. Finally, the academically disadvantaged student population has a higher percentage of overage students, i.e. students that are at least a year older than their grade level peers (45%) than the full student population (21%). The statistics by subject indicate the proportion of observations in each of the subjects included in this analysis. Disadvantaged students are over represented in Algebra I, English I, and Biology compared to the overall sample as they are less likely to take higher level EOC tested courses compared to other students.

Table A.1: Student Characteristics

		Full Sample N= 866,790	**Disadvantaged Student Sample N = 75,313**
		Mean (sd)	**Mean (sd)**
EOC Score Average (Non-DSSF)		60.4 (8.96)	51.1 (6.41)
EOC Score Average (DSSF)		56.7 (7.95)	50.7 (6.15)
Standardized EOC Score Average (Non-DSSF)		-0.21 (0.98)	-1.06 (0.73)
Standardized EOC Score Average (DSSF)		-0.44 (0.89)	-1.11 (0.70)
Grade 8 reading score (N =792,048)		0.17 (0.90)	-1.13 (0.73)
Grade 8 mathematics score (N=793,134)		0.17 (0.94)	-1.18 (0.51)
Gender: Male (N = 856,134)		49%	50%
Days absent		7.84 (9.30)	10.80(11.42)
Class Characteristics	Peer ability	0.16 (0.62)	-0.41 (0.46)
	Advanced curriculum	26%	3%
	Remedial curriculum	1%	1%
Race	White	63%	33%
	Black	28%	56%
	Asian	2%	1%
	Hispanic	4%	6%
	Multi-racial	2%	1%
	American Indian	1%	2%
Income	Free lunch	20%	47%
	Reduced lunch	5%	8%
	Missing	9%	<1%
Parent Education	< High school	6%	13%
	HS degree	20%	35%
	Some college	30%	29%
	College	34%	22%
	Missing	10%	1%

Table A.1: Student Characteristics cont..

		Full Sample N= 866,790	Disadvantaged Student Sample N = 75,313
		Mean (sd)	**Mean (sd)**
Exceptionality	Gifted	13%	0%
	High incidence	5%	15%
	Cognitive	6%	3%
	Behavioral	6%	2%
	Sensory	4%	0%
	Physical	11%	5%
	Severe	9%	2%
Limited English Proficiency	Previous LEP	5%	0%
	Current LEP	9%	3%
Grade level	Underage	1%	1%
	Overage	21%	45%
	Grade 9	46%	41%
	Grade 10	31%	29%
	Grade 11	18%	32%
	Grade12	5%	9%
End of Course tests	English 1	0.19 (0.39)	0.25 (0.43)
	Algebra 1	0.21 (0.41)	0.27 (0.45)
	Algebra 2	0.12 (0.33)	0.06 (0.24)
	Geometry	0.13 (0.33)	0.07 (0.26)
	Physical Science	0.08 (0.27)	0.13 (0.34)
	Biology	0.16 (0.37)	0.18 (0.39)
	Chemistry	0.08 (0.27)	0.02 (0.15)
	Physics	0.02 (0.13)	0.00 (0.03)

2004-05 and 2005-06 Classroom and Teacher Data

The classroom and teacher data were provided by NC DPI from administrative records. These data include information on the specific students assigned to each classroom, the course title, and the instructor's name. Since all students within a class can be identified with these data, it is possible to calculate the peer ability dispersion value – which represents the range of abilities within each classroom. A value of zero indicates that all students within the class had the same average score on their reading and math 8^{th} grade EOG assessments. Using the course titles, the research team created indicators for advanced and remedial curriculum classrooms. Courses containing the words or phrases "honors", "advanced", or "ap" were coded as advanced curriculum courses. Those containing "remedial" or "special" were coded as remedial courses. The instructor's name provided a means to link teacher quality indicators and salary information to student level data. The data also provide information on the classroom racial and income composition of the students used in descriptive tables on the following pages.

Table A.2 shows there is a slightly wider within class peer dispersion for academically disadvantaged students. Also, not surprisingly, a significantly smaller percentage of academically disadvantaged students enrolled in classes that use an advanced curriculum (6% v. 23% of the full sample). Reflecting the individual level data, the classroom composition data for academically disadvantaged students have a higher percentage of minorities and free lunch students in their classrooms.

The overall teacher quality data were provided by the NC DPI from administrative records. The 2004-05 and 2005-06 data from DPI include information on teacher quality in DSSF districts as well as teacher quality in the other districts in North Carolina. The DPI data measure many of the teacher quality variables that research suggests are important. Taken as a whole, the data give us a reasonable set of indicators of teacher quality, both in the DSSF districts and statewide.

Research generally supports the proposition that teacher quality shapes student outcomes (Ferguson, 1998; Hanushek, Kain, and Rivkin, 1999; Jordan, Mendro, and Weersinghe, 1997; Sanders and Rivers, 1996; Wright, Horn, and Sanders, 1997). Research is not always unequivocal about which specific dimensions of teacher quality actually make a difference in student outcomes, but it is reasonably clear that certain variables are likely to matter. Teachers scoring higher on tests of verbal ability and teachers with higher general academic proficiency appear to be more effective in promoting student learning, presumably because they are better at presenting material clearly and helping students understand course content (Bowles and Levin, 1968; Coleman et al., 1966; Hanushek, 1971; Murnane, 1985; Strauss and Vogt, 2001). Therefore we include a variable (mean test score) that is the normed average of the individual teachers most recent standardized test scores, including Praxis, NTE, and SAT exams.

Simply holding an advanced degree may not improve teacher and student performance, but advanced education in the subject that the teachers actually teaches does seem to increase teachers' effectiveness (Hawk, Coble, and Swanson, 1985). Coursework in teaching and learning also seems to help (Ferguson and Womack, 1993), especially in combination with solid subject matter knowledge (Druva and Anderson, 1983). Formal teacher preparation and licensure requirements seem to help assure that teachers have both of these (Darling-Hammond, Holtzman, and Heilig, 2005). Inexperienced teachers are generally less effective than teachers with three to five years or more of experience (Klitgaard and Hall, 1974; Murnane and Phillips, 1981; Rosenholtz, 1986). However, after about twenty-five years in the classroom, many teachers' effectiveness seems to decline. Not surprisingly, the evidence seems strongest for teachers with a combination of strong subject matter knowledge, knowledge of teaching and learning, and several years of experience (Ferguson, 1991; Ferguson and Ladd, 1996). Some research indicates that National Board Certified Teachers are more effective than their uncertified counterparts (Goldhaber and Anthony, 2004). To look at each of these factors, we include variables related to years experience, NBCT, infield teaching, and if the teacher has an advanced degree.

For this report, we linked teacher characteristics to their classroom assignments and created variables for each teacher as follows: (a) an indicator variable equal to one for teachers with a bachelor's degrees from a "highly competitive" or "most competitive" institution of higher education[4], an indirect measure of their general academic proficiency; (b) an indicator variable equal to one for teachers with Master's or Doctoral degrees; (c) the general academic ability of each teacher using the teacher's average score on teacher preparation exams such as the PRAXIS, NTE, or GRE; (d) an indicator variable equal to one if the teacher was teaching a class for which they held a continuing or initial license in the field; and (e) an indicator variable indicating the teacher's level of teaching experience.

Finally, teacher salary characteristics were provided by expenditure files from NC DPI for 2004-05 and 2005-06. A teacher's regular pay included all the expenditures made from purpose codes 5100, 5200, 5500 and object codes 121, 123, 125, 126, 128, and 129. Any individual who received a payment from any of these purpose and object codes was considered a teacher. The 2004 NC Chart of Account's definitions of the object codes are described below:

NC Chart of Accounts for 2004: Teacher Salary

Expenditure Object Code	NC Chart of Accounts 2004 Description
121 – Teacher	Include the salary of the person assigned to instruct pupils not classified elsewhere: (i.e., academic instruction, vocational education, library/media specialist, and guidance services).
123 – Teacher – Other	Include the salary of the person assigned to teach ROTC, the salary of the person assigned as lead teacher in the summer school program, and/or other state designated purposes. This special designation is required to permit proper identification for salary assignment purposes by DPI.
125 – New Teacher Orientation	Include the salary of the person attending assigned new teacher orientation, outside of the teacher's contract calendar, not to exceed three days.
126 – Extended Contract Days	Include pay to teachers at designated schools for additional contract days beyond the school year.
127 – Interim Teacher – Non-certified	An interim employee may be employed when a vacancy in a teaching position occurs. Include the salary of a person being paid at a non-certified rate (substitute pay rate). The interim employee is not licensed in the area of assignment. (See State Salary Manual, Section D.II.O.)
128 – Retired Teacher – Exempt from the Earnings Cap	Include the salary of retired teachers who have not been employed "in any capacity with a public school, except as a substitute teacher, for at least 6 months immediately preceding the effective date of reemployment". They shall not be subject to the computation of post retirement earnings (earning cap). (See Benefits Manual, Section 16.)
129 – Other Professional Educator Assignments	Include the salary of the person(s) assigned to perform professional educator activities, which include but are not limited to, Visiting International Faculty (VIF) and Vocational Competency Achievement Tracking System (VoCats).

[4] These rankings are based on institutional rankings provided by ***Barron's Guide to the Most Competitive Colleges* (2005)**. Barron's ranks degree granting intuitions on a scale from 1 (not competitive) to 6 (most competitive). Rankings are primarily based on acceptance rates and college entrance examination scores. For this variable, we have collapsed the top two rankings (highly competitive and most competitive) into one variable.

Supplementary Pay

If an individual was identified to be a teacher, their supplementary pay was obtained by adding any additional payment received under object codes 181, 183, 184, and 187. The 2004 NC Chart of Account's definitions of these object codes are described below:

NC Chart of Accounts for 2004: Supplemental Pay

Expenditure Object Code	NC Chart of Accounts 2004 Description
181 - Supplementary Pay	Include supplements paid to employees from local, federal, and/or certain State funds that are determined to be amounts in addition to salary paid for the individual.
183 - Bonus Pay	Include legislated bonus payments made to eligible employees.
184 - Full-Time Substitute	Salary of the person employed for at least 30 hours per week, and who is expected to be employed full-time for at least six (6) consecutive months as a substitute for a teacher, who is on paid leave.
187 - Salary Differential	Include the salary differential amount paid to employees from local, federal and/or certain State funds that are determined to be a part of the salary paid for the individual. Examples: (1) ROTC differential – paid in addition to the State certified salary; (2) Military differential (PRC 021) – paid to active duty military the difference between the State salary and the active duty pay.

Two teacher salary variables were calculated using the NC DPI administrative data: total non-bonus pay and total bonus pay. As shown in Table A.2, teachers across all classrooms make slightly more than teachers of disadvantaged students.

Table A.2: Classroom and Teacher Characteristics

DSSF		All Classrooms N= 46,127	Classrooms with Disadvantaged Students N = 25,493
		Mean (sd)	Mean (sd)
Class Characteristics	Number of students	18.79 (6.54)	18.56 (6.16)
	Peer ability dispersion	0.59 (0.15)	0.63 (0.15)
	Advanced curriculum	23%	6%
	Remedial curriculum	1%	1%
Race Compositions	White	61% (0.21)	52% (0.29)
	Black	29% (0.27)	38% (0.27)
	Asian	1% (0.05)	1% (0.04)
	Hispanic	4% (0.08)	5% (0.09)
	Multi-racial	2% (0.04)	2% (0.04)
	American Indian	1% (0.07)	2% (0.08)
Income Averages	Free lunch	24% (0.20)	31% (0.20)
	Reduced lunch	6% (0.08)	7% (0.08)
Teacher quality	NBC	12% (0.32)	9% (0.29)
	Infield certification	83% (0.37)	80% (0.40)
	BA degree	100%	100%
	Advanced degree	32% (0.47)	29% (0.46)
	Most/ high competitive	19% (0.40)	18% (0.38)
	Mean test score	0.15 (0.69)	0.09 (0.68)
Teacher salary characteristics (in 1,000s)			
	Salary non-bonus	34.31 (8.95)	33.54 (8.77)
	Total Bonus	2.18 (2.26)	2.18 (2.24)
Teacher Experience	First year	6% (0.24)	7% (0.26)
	1 – 2 years	11% (0.31)	12% (0.33)
	3-4 years	9% (0.28)	10% (0.30)
	5-9 years	22% (0.42)	23% (0.42)
	10-14 years	15% (0.36)	14% (0.35)
	15-19 years	11% (0.32)	10% (0.30)
	20-24 years	9% (0.28)	8% (0.27)
	25 plus years	17% (0.38)	16% (0.36)

2004-05 and 2005-06 School Data

Finally, NC DPI provided school level data from administrative school records. For our purposes, a high school was defined as a regular non-charter school that contained at least 9th through 12th grade and was not a school designed to serve students with special needs such as an alternative school or a school serving students with a specific disability. The primary variable of interest for our analysis is the DSSF variable. Approximately 9% of students are in a DSSF high school, while 10% of academically disadvantaged students are in a DSSF high school. As shown in Table A.3, at the school level, the composition measures do not vary substantially between the full sample of students and just the academically disadvantaged students.

Using NC DPI 2004-05 and 2005-06 expenditure data, per pupil spending was categorized into eight different categories ranging from regular instruction expenses to school leadership. Also included are expenditure variables for the amount of DSSF dollars spent per pupil, and the LEA level of per pupil expenditures. A more detailed discussion of per pupil expenditures is below.

Table A.3: School Characteristics

		All Schools N= 330	**School Disadvantaged Student Sample N = 311**
		Mean (sd)	**Mean (sd)**
School Characteristics	Number of students	1,130 (514)	1,141 (515)
	DSSF	9%	10%
	Combined characteristics	0.18 (0.14)	0.18 (0.14)
Race Compositions	White	61% (0.25)	61% (0.25)
	Black	30% (0.23)	30% (0.23)
	Asian	2% (0.02)	2% (0.02)
	Hispanic	5% (0.04)	5% (0.04)
	Multi-racial	2% (0.01)	2% (0.01)
	American Indian	1% (0.06)	1% (0.06)
Income Averages	Free lunch	25% (0.13)	25% (0.13)
	Reduced lunch	7% (0.04)	7% (0.04)
Per Pupil Expenditures	DSSF per pupil	$230 (77)	$240 (77)
	Regular instruction	$3380 (687)	$3369 (574)
	Special instruction	$616 (310)	$620 (312)
	Professional development	$37 (28)	$36 (27)
	Supplemental education services outside school day	$66 (69)	$67 (69)
	Student Services	$317 (137)	$314 (133)
	Technology	$30 (41)	$30 (41)
	School leadership	$412 (165)	$418 (160)
	LEA per pupil instruction	$1,419 (606)	$1,414 (613)

Methods

Now that we have provided a description of the data that were used in the analysis, we turn to a discussion of the analysis methods used. This discussion will begin with matching individual students to their classroom teachers, followed by a discussion of how districts were selected into the DSSF program, and the analysis modeling technique (HLM).

Roster Matching

By carefully combining several types of data supplied by DPI, we were able to connect individual teachers in specific EOC subject courses with individual students. Table A.4 provides descriptive statistics on the success of data matching. About 90% of all roster observations were able to be matched to a teacher. Once students' prior test histories were integrated, about 72.5% of the total number of roster entries linked a roster entry with a teacher identifier and a student's prior 8th grade test scores. A little over 80% of observations linked to teachers were able to be matched with a student's 8th grade test scores. With these data, we were able to not only contrast the quality of teachers teaching tested subjects across DSSF districts and other districts, we were able to control for the effect of different levels of teacher quality indicators on student outcomes. Thus, at the high school level, for example, we could tell which teachers taught Algebra I to which students, not only within DSSF districts but statewide. This matching allows us to compare the achievement of students in DSSF districts with students in the rest of the state, and to correctly assess the extent to which individual student characteristics, the characteristics of their classroom teachers, and school environmental characteristics impact their achievement. This type of matching allows more precise estimates of the impact of the program on student performance, since we can adjust for differences in classroom and school characteristics that might not relate to DSSF.

Table A.4: Matching Statistics - End of Grade Tested Subjects in Math, English, and Science in 2004-05 and 2005-06

	Number of Cases	Percent Matched (All Observations)	Percent Matched (Of Teacher Matches)
Unique Observations in Roster File (All Observations)	1,194,849.00	---	---
Observations Matched to Teachers	1,076,066.00	90.1%	---
Observations Matched to Teacher and Students' Prior Test Scores	866,790.00	72.5%	80.6%

Selection into the DSSF Program

DSSF districts were selected based on a combined score from four variables: proportion of proficient students, teacher retention, percent of teachers with more than five years teaching experience, and percent living above the poverty line. From that list of characteristics, each district received a score, labeled 'combined characteristics'. Districts with the greatest educational need according to this formula had the lowest scores. Therefore, in order to compare the DSSF districts with the remaining districts we employed a regression discontinuity (RD) design. In selecting the school districts to receive DSSF funding, the 16 lowest scoring districts on the combined characteristics variable received funding. Because we know the criteria used for selection into the DSSF program, we can estimate the unbiased effects of DSSF on student achievement (Shadish, Cook and Campbell 2002; Trochim, Cappelleri, and Reichardt 1991; Van der Klaauw 2002)**.**

Approach to the Analysis: Hierarchical Linear Modeling

All regression analyses presented within this report rely on a statistical approach known as Hierarchical Linear Modeling (HLM). This type of regression analysis is specifically designed to analyze data that are nested in which the context of the environment may impact the outcome (in this case, students within classes within schools). Looking at schools as an example, we generally observe that learning environments vary from school to school and are influenced by the types of expenditures made within schools and variations in the socio-demographic characteristics of schools. The HLM approach provides a framework to correctly analyze individual observations nested within schools and classrooms. The non-random nesting of data occurring within educational context violates a key assumption of other types of regression known as ordinary least squares regression, that is individual observations are independent of one another. Observations cannot be completely independent because they occur within classrooms, and schools, where individual variation within these contexts is to some extent explained by grouping. A simple statistical test determines the appropriateness of HLM for analyzing data. The intraclass correlation coefficient (ICC) expresses the proportion of variation occurring with each level of a nested data context (Raudenbush and Bryk, 2002). In cases where the proportion of variation attributable to grouping is greater than about 10 percent, HLM analysis is appropriate (ibid). For the full sample of students, the ICC was 11.1% for school level grouping and 34.7% for classroom level grouping meaning that up to 11.1% of differences in student scores could potentially be explained by school characteristics and an additional 34.7% by classroom characteristics. The ICC is lower for the sample that included only disadvantaged students, but was well over the minimum threshold required to indicate that HLM analysis is appropriate.

Net Effects Model Approach

As a first step in determining the impact of DSSF on EOC student achievement, we ran a net effects model. This model provides an estimate of the effect of DSSF funding on student achievement leaving out variables that might have been affected by the DSSF pilot such as teacher quality, class size and per pupil expenditures. Because DSSF provides money to support teacher bonuses which might result in higher quality teachers, technology, and other types of expenditures, positive effects of DSSF could be hidden by the inclusion of variables whose values changed due to receipt of DSSF funds. The traditional way to obtain an unbiased effect of DSSF using the regression discontinuity design is to include a single variable whose value is one for DSSF schools and zero for non-DSSF schools along with the assignment variable (educational need) and student and school controls that could not be affected by DSSF. The DSSF indicator variable estimates the average difference in student test scores for students enrolled in schools in the pilot districts, holding constant the student and school controls included in models. One can interpret the value as the average difference in achievement between two students enrolled in the same type of course with the same personal, classroom, and school characteristics where the only difference is enrollment in a DSSF school. For a robustness check, not included in the report, the DSSF indicator variable was replaced with a continuous variable indicating the dollar amount of DSSF funded per pupil expenditures made within a school. All non-DSSF schools have values of zero, while DSSF schools average values of about $250. Both models control for student characteristics, classroom characteristics, and school characteristics that are not influenced by the DSSF program.

Within this net effects framework and holding constant differences between students, classrooms, and schools on a number of measures unrelated to DSSF we find a positive and statistically significant impact of being in a school provided with DSSF funding. Students within DSSF schools average EOC test scores about 0.16 standard deviations higher than comparable students. This difference persists for a variety of functional forms of the educational need variable including the combined measure, its square, and its cube. The alternative model which treats DSSF expenditures as a continuous variable is also positive and statistically significant. The size of the effect, about 0.05 standard deviations indicates that an additional $100 of per pupil DSSF expenditures per pupil increases the average student EOC test score by about one twentieth of a standard deviation (0.05). The average per pupil DSSF expenditure within schools receiving DSSF funding was about $150 resulting in an average impact of about 0.0795 standard deviations (0.053 * 1.5). The magnitude of this effect is about half the size of the model utilizing an indicator variable for schools receiving DSSF funding. Additionally, local education agencies spent an additional $100 per pupil in DSSF funds. Assuming a similar effect on student test score outcomes from these expenditures, we would expect a total effect of about 0.1325 (0.0795 + 0.053) similar to the value observed in the regression analysis using the indicator variable.

The net effects equation is:

$$\begin{aligned} Y_{ics} \\ = \gamma_{000} + \gamma_{001} DSSF_{Indicator} + \gamma_{002} EdNeed + \gamma_{003} SchoolSize + \gamma_{004} PropByRace + \gamma_{005} Year \\ + \gamma_{006} PropFRLunch + \gamma_{010} AdvCurriculum + \gamma_{020} RemedialCurriculum \\ + \gamma_{030} ClassroomAbilityDispersion + \gamma_{100} G8ReadScore + \gamma_{200} G8MathScore + \gamma_{300} Absenses \\ + \gamma_{400} PeerAbility + \gamma_{500} Gender + \gamma_{600} Race + \gamma_{700} FreeReducedLunchStatus + \gamma_{800} Exceptionality \\ + \gamma_{900} LEP + \gamma_{1000} ParentalEducation + \gamma_{1100} Age + \gamma_{1200} Grade + \gamma_{1300} TestType + u_{00s} + r_{0cs} + e_{ics} \end{aligned}$$

Example of mediated effects model for class size as a mediator of DSSF program effects:

$$\begin{aligned} Y_{ics} \\ = \gamma_{000} + \boldsymbol{\gamma_{040} ClassSize} + \gamma_{001} DSSF_{Indicator} + \gamma_{002} EdNeed + \gamma_{003} SchoolSize + \gamma_{004} PropByRace \\ + \gamma_{005} Year + \gamma_{006} PropFRLunch + \gamma_{010} AdvCurriculum + \gamma_{020} RemedialCurriculum \\ + \gamma_{030} ClassroomAbilityDispersion + \gamma_{100} G8ReadScore + \gamma_{200} G8MathScore + \gamma_{300} Absenses \\ + \gamma_{400} PeerAbility + \gamma_{500} Gender + \gamma_{600} Race + \gamma_{700} FreeReducedLunchStatus + \gamma_{800} Exceptionality \\ + \gamma_{900} LEP + \gamma_{1000} ParentalEducation + \gamma_{1100} Age + \gamma_{1200} Grade + \gamma_{1300} TestType + u_{00s} + r_{0cs} + e_{ics} \end{aligned}$$

The complete list of variables corresponding to each of the categories above can be found in Tables 1a and 1b.

References

Bowles, S., & Levin, H. M. (1968). The Determinants of Scholastic Achievement--An Appraisal of Some Recent Evidence. *Journal of Human Resources, 3*(3), 3-24.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weindfeld, F. D., et al. (1966). *Equality of Educational Opportunity*. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare, Office of Education.

Darling-Hammond, L, Holtzman, D.J., Gatlin, S.J.,& Heilig, J. V. (2005). Does Teacher Preparation Matter? Evidence about Teacher Certification, Teach for America, and Teacher Effectiveness. *Education Policy Analysis Archives, 13*(42).

Druva, C. A., & Anderson, R. D. (1983). Science Teacher Characteristics by Teacher Behavior and by Student Outcome: A Meta-Analysis of Research. *Journal of Research in Science Teaching, 20*(5), 467.

Ferguson, P., & Womack, S. T. (1993). The Impact of Subject Matter and Education Coursework on Teaching Performance. *Journal of Teacher Education, 44*(1), 55.

Ferguson, R. F. (1991). Paying for Public Education: New Evidence on How and Why Money Matters. *Harvard Journal on Legislation, 28*(2), 465-498.

Ferguson, R. F. (1998). “Can Schools Narrow the Black-White Test Score Gap?” In C. J. a. M. Phillips (Ed.), *The Black White Test Score Gap* (pp. 318-374). Washington, D.C.: Brookings Institution Press.

Ferguson, R. F., & Ladd, H. F. (1996). How and Why Money Matters: An Analysis of Alabama Schools. In *Holding Schools Accountable: Performance-based Reform in Education* (pp. 92-265): Washington, D.C. Brookings Institution.

Goldhaber, D. & Anthony, E., (2004). Can Teacher Quality be Effectively Assessed?, from http://www.ubrban.org/url.cfm?ID=410958

Hanushek, E. (1971). *Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data*: American Economic Review.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). *Do Higher Salaries Buy Better Teachers?* Working paper 7062. Cambridge , MA: National Bureau of Economic Research.

Hawk, P. P., Coble, C. R., & Swanson, M. (1985). Certification: It Does Matter. *Journal of Teacher Education, 36*(3), 13-15.

Henry, G., Thompson, C., & associates (2008) *Final Report of the High School Resource Allocation Study.* Chapel Hill, NC: Carolina Institute for Public Policy.

Jordan, H. R., Mendro, R. L., & Weersinghe, D. (1997). *Teacher Effects on Longitudinal Student Achievement: A Preliminary Report on Research on Teacher Effectiveness* Paper presented at the National Evaluation Institute, Indianapolis, IN. Kalamazoo, MI. 93

Klitgaard, R., & Hall, G. (1974). Are There Unusually Effective Schools? *Journal of Human Resources, 74*, 90-106.

Murnane, R. J. (1985). *Do Effective Teachers Have Common Characteristics: Interpreting the Quantitative Research Evidence*. Paper presented at the National Research Council Conference on Teacher Quality in Science and Mathematics.

Murnane, R. J., & Phillips, B. R. (1981). Learning by Doing, Vintage, and Selection: Three Pieces of the Puzzle Relating Teaching Experience and Teaching Performance. *Economics of Education Review, 1*(4), 453.

Raudenbush, S. & Bryk, A. (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed. Thousand Oaks: CA.

Sanders, W. L. & Rivers, J.C. (1996). Cumulative and Residual Effects of Teachers on Future Student Academic. *University of Tennessee Value-Added*, from http://www.heartland.org/pdf/21803a.pdf

Shadish, W., Cook, T., & Campbell, D. (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal inference*. Boston: Houghton Mifflin.

Strauss, R., & Vogt, W. (2001). *It's What You Know, Not How You Learned to Teach It: Evidence from a study of the effects of knowledge and pedagogy on student achievement.* Paper presented at the American Educational Finance Association, Cincinnati.

Trochim, W., Cappelleri, J., & Reichardt, C. (1991) Random Measurement Error Does Not Bias the Treatment Effect Estimate in the Regression-Discontinuity Design II. When an Interaction Effect Is Present. *Evaluation Review*, Vol. 15, 571-604.

van der Klaauw, W. (2002) Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression- Discontinuity Approach. *International Economic Review*, Vol. 43, 1249-1287.

Wright, S. P., Horn, S. P. & Sanders, W. L. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67.

Carolina Institute for Public Policy

UNC
COLLEGE OF
ARTS & SCIENCES