Targeted Funding for Educationally Disadvantaged Students: A Regression Discontinuity Estimate of the Impact on High School Student Achievement
Author(s): Gary T. Henry, C. Kevin Fortner and Charles L. Thompson
Published by: American Educational Research Association
Stable URL: http://www.jstor.org/stable/40732417
Accessed: 12-01-2016 17:45 UTC
Targeted Funding for Educationally Disadvantaged Students: A Regression Discontinuity Estimate of the Impact on High School Student Achievement

Gary T. Henry
C. Kevin Fortner
University of North Carolina at Chapel Hill

Charles L. Thompson
East Carolina University

Evaluating the impacts of public school funding on student achievement has been an important objective for informing education policymaking but fraught with data and methodological limitations. Findings from prior research have been mixed at best, leaving policymakers with little advice about the benefits of allocating public resources to schools or how it might best be done. In this study, the authors take advantage of a pilot supplemental funding program in North Carolina that used a quantitative index of educational advantage to select the most educationally disadvantaged districts in the state to receive funding. The targeted districts received supplemental funds of $250 per pupil or $840 per academically disadvantaged pupil for the 2 years of the pilot. Using a regression discontinuity design and multilevel models with extensive controls, the authors estimate that the marginal average treatment effect of the supplemental funding was 0.133 standard deviation units and that the effect on educationally disadvantaged students was 0.098 standard deviation units. The treatment effect represents approximately one third of the difference between the average score in top performing and low performing high schools.

Keywords: disadvantaged students, education funding, high schools, regression discontinuity

Whether providing additional public funds for schools produces educational benefits for students is an often assessed but still unresolved policy question. Hanushek (1997) identified no less than 377 estimates of the relationship between educational resources and performance, drawn from 90 published studies, but the results were largely mixed concerning the effects of more funding on student achievement. In spite of these studies and many others that have been conducted since 1997, no consensus has emerged from the evidence estimating the impact of educational expenditures on student performance (see, e.g., Greenwald, Hedges, & Laine, 1996; Hanushek, 1997, 2003; National Research Council, 1999; Taylor, 2001).

The lack of clarity regarding the relationship between resources and achievement may...
be attributable to a number of factors, including data limitations, the correlational designs that have been typically employed, and the practice of making undifferentiated expenditure changes across entire states. Historically, both school finance and student achievement data have been aggregated to the district and sometimes the state level, potentially masking important variation within districts and within schools, which may have inflated the estimated effects (Hanushek, Rivkin, & Taylor, 1996). It is also possible that limitations in the data and the analytical techniques have attenuated the effects of certain additional resources, such as class size reduction, that the expenditures were used to fund (Boozer & Rouse, 2001). In most cases, studies of the effects of resources on student achievement have relied on a specification of an education production function (EPF) using regression techniques to estimate the effect of resources on test scores, including level of achievement, value added models, and gain scores. These estimates are distinctly different from estimates derived from experimental situations (Todd & Wolpin, 2003). The estimates from EPF analyses are ceteris paribus estimates, or estimates net of other controls, which means that these methods control for variations in other inputs that may occur in response to the treatment in an experimental situation. For example, if additional resources for teachers’ salaries affect decisions about deploying technology within schools, these effects are captured in experimental estimates of the effects of teachers’ salaries on student test scores but not in EPF estimates if technological inputs are included. In addition, EPF estimates of the effects of specific inputs or resources may vary when alternative specifications are employed, and the estimates are subject to omitted variable bias when measures of important household resources or student endowments are not included or are measured with error (Barrow & Rouse, 2005; Hanushek, Kain, Markman, & Rivkin, 2003; Todd & Wolpin, 2003). Recently, Dee and Levine (2004) and Guryan (2003) have exploited significant year-to-year funding increases from a state finance reform that amounted to approximately $475 per pupil (Dee & Levine, 2004) using a difference-in-differences estimation to suggest that additional state funds for education did increase district spending and did have an effect on educational resources and achievement.

In this study, we continue the progress toward incorporating more rigorous design elements in an effort to accurately estimate the effect of increased state funding on student achievement. To produce consistent or, in Rubin’s (1974, 2005) terms, unbiased treatment effect estimates, there is a growing consensus that alternatives to correlational designs are preferred. A strong preference has emerged for estimating effects from designs employing random assignment to treatment and control, which require the fewest number of assumptions for causal assignment and, in theory, distribute confounds equivalently across treatment and controls (Cook, Shadish, & Wong, 2008; Glazerman, Levy, & Myers, 2003; Rubin, 1974, 2005; Wilde & Hollister, 2007). In the absence of randomized experiments or prior to the decision to allocate resources for a randomized experiment, regression discontinuity (RD) designs using proper specifications of the estimating equations have been shown to yield unbiased (consistent) and valid estimates of program effects (Cook et al., 2008; Goldberger, 1972; Imbens & Lemieux, 2008; Schochet, 2008; Shadish, Cook, & Campbell, 2001; van der Klaauw, 2008). These designs have been employed in the evaluation of several education programs, including the effects of remedial education (Jacob & Lefgren, 2004) and prekindergarten programs (Gormley & Gayer, 2005; Gormley, Phillips, & Gayer, 2008). Other promising designs, including propensity score matching (Cook et al., 2008; Diaz & Handa, 2005; Rosenbaum & Rubin, 1983), have been developed and used to estimate the effects of educational programs (Henry, Gordon, & Rickman, 2006). In this case, the RD design was selected because the strong theory shows that this design is capable of producing unbiased estimates of effects (Goldberger, 1972) with few and weaker assumptions for making causal inferences than all other designs with the exception of random assignment studies.

RD designs require that scores on a quantitative variable determine assignment to treatment to yield unbiased estimates of the marginal average treatment effects (MATEs). Because of the fact that the assignment variable is fully known, the selection process is fully known and its...
impact can be controlled. Imbens and Lemieux (2008) note that the unbiased estimate of an effect depends on a smooth functional relationship between the assignment variable and the outcome variable. In practice, meeting this assumption requires that the functional form of the assignment variable is correctly specified in the estimating equations. The inclusion of the proper functional form of assignment variable in the estimation equations “soaks up” the correlation with the disturbance term, thereby eliminating bias in the estimate of the coefficient on the variable indicating assignment to treatment. In addition to the proper specification of the equation, the credibility of the effect estimates depends on three assumptions that we assess: (a) The treatment occurred and was sharply discontinuous at the cutoff, (b) mechanisms were activated by the treatment that plausibly link the treatment to the effect, and (c) there were no “hidden” treatments or confounding covariates that were similarly discontinuous at the cutoff (Reichardt & Henry, in press).

The use of a RD design to assess the effect of additional funding on test scores was made possible when school districts in North Carolina were assigned to the Disadvantaged Student Supplemental Fund (DSSF) pilot program based strictly on an index of educational advantage. The 16 districts with the lowest scores on the educational advantage index were selected for the pilot program in the summer of 2004. Additional funding of $250 per pupil, which averaged $840 per academically disadvantaged pupil, was allocated by the state to these districts each year for 2 years during the pilot phase of the program. For estimating the impact of additional resources on students in educationally disadvantaged school districts, the situation is nearly ideal. However, extrapolating the effect to other districts must be approached cautiously, if warranted at all. Nonetheless, an unbiased estimate of MATE for targeted supplemental funding has great relevance to educational policy discussions about the means for improving educational performance in districts plagued by educational disadvantages and underachievement. It is important to note that the ability to generate unbiased estimates is not dependent on the ability to perfectly or even adequately measure educational advantage; it is dependent on the assignment to treatment using a strict cutoff on the quantitative measure of educational advantage as well as the validity of the assumptions mentioned above. Of course, the accuracy of the targeting would be affected by the validity of the measure of educational advantage and should be a concern for policymakers seeking to reduce underachievement.

To preview the findings, we estimate that the MATE of the supplemental funding on high school test scores was 0.13 standard deviation units and that the MATE on a subpopulation of educationally disadvantaged students was 0.098 standard deviation units. These estimates were produced using mixed models with individual student end-of-course (EOC) tests as the outcome measure and extensive controls for the individual characteristics, including prior test scores, family characteristics, and school composition. To explain the program and assess the plausibility of the assumptions underlying the estimation of unbiased effects from the RD design as well as the study methods, we proceed as follows: Section 2 describes the program; Section 3 lays out the evidence that the treatment did in fact occur and was sustained for the 2-year study period that the other reforms taking place within the state were not discontinuous around the cutoff for supplemental funding and the linkage between the treatment and the effects; Section 4 describes the study design, sample, measures, and analytical model; Section 5 explains the findings; and Section 6 provides our conclusions concerning the findings and their implications for policy.

2. The North Carolina DSSF

In 2004, Governor Mike Easley and the North Carolina Board of Education initiated the DSSF to improve the education of academically at-risk students. The pilot provided $22.4 million to fund improvements in 16 school districts in North Carolina during 2004–2005. The 16 districts selected for the pilot were the state’s most disadvantaged as indicated by their scores on an index of educational advantage. The index of educational advantage was developed by the governor’s office and the staff of the state board of education prior to the announcement of the DSSF. Developing the index and funding the
program from unspent state funds in the confines of the executive branch prior to announcing the program to the legislature, school boards, educators, and the public virtually precluded any manipulation of the index or the values assigned to individual school districts, which has been shown to be a problem for obtaining unbiased effect estimates from RD designs (McCrary, 2007; Reichardt & Henry, in press). The index included four (equally weighted) measures of educational advantage: teacher stability, experienced teachers, children not living in poverty, and students meeting state proficiency standards. The measures were recorded in percentage terms and summed to produce the index of educational advantage. The pilot continued in these 16 districts in 2005–2006 with a slight funding increase. In 2006–2007, the 16 pilot districts received the same per pupil amount, which ranged from $264,500 to $6.1 million. In addition, the state’s 99 other districts received approximately $24 million, or $88 per disadvantaged pupil, which is 10.4% of the amount per disadvantaged pupil in the pilot districts. The current study assesses the impact of the first 2 years of pilot funding.

The overarching goal of the DSSF pilot was to improve student learning and performance, especially for educationally disadvantaged students. For the purpose of this evaluation, disadvantaged students are students currently enrolled in high school who did not achieve proficiency on the state’s assessments for either reading or math the year before they entered high school (eighth grade). The DSSF pilot allowed school districts flexibility to spend funds on choices from a “menu of proven strategies.” The menu included a dozen distinct strategies designed to (a) improve pilot districts’ competitive position in recruiting and retaining teachers via signing, performance, and retention bonuses, (b) strengthen teachers’ skills via mentorship for new teachers as well as professional development for teachers more broadly, (c) reduce class sizes, (d) upgrade materials and equipment, (e) promote the use of annual assessment data to guide assignment, compensation, and selective retention of teachers, (f) support the use of interim assessment data to guide instruction, (g) extend instructional time, (h) diagnose individual disadvantaged students’ needs and develop individual education plans for them, and (i) provide classroom support for limited English proficiency (LEP) students. Each district was given flexibility to decide how to allocate funds to meet its particular needs. Thus, although the allocation of supplemental funds was $250 per pupil from the state to the 16 pilot districts, the amounts allocated to the schools within the district varied slightly, depending on how the funds were used within the district. Plans for allocating the supplemental resources were submitted and approved by local school boards and by the North Carolina Department of Public Instruction (NCDPI). The NCDPI was required to provide technical assistance to districts and to monitor the implementation of the intervention.

3. Can the Additional State Funds Be Plausibly Linked to Effects on Student Test Scores?

The hypothesis that we test in this study is that supplemental state funding targeted to educationally disadvantaged districts for the purpose of improving performance, especially for academically disadvantaged students, can actually improve test scores. The funding was not a “one-size-fits-all” approach for a prescribed set of uniform program activities for each of the districts but an increase in funding that could be tailored to meet the needs of the particular schools in the pilot districts. Of course, it is always possible that the districts could return all or part of the supplemental funding to the local taxpayers rather than exhibiting the “flypaper effect,” which results in spending more money for education (Dee & Levine, 2004; Hines & Thaler, 1995).

In this section, we document that dollars provided through the DSSF produced greater increases in overall per pupil expenditures than those in all non-DSSF North Carolina districts and in the 16 districts closest to DSSF pilot districts on the index of educational advantage used to select the pilot districts. Furthermore, the amount of the increases in overall per pupil expenditure in the DSSF districts mirrors the magnitude of the supplemental funds supplied by the state, indicating that pilot districts did not use DSSF funds to supplant local funds. Thus, the DSSF did in fact represent a real treatment.
Observing a sharp discontinuity in the funding increases between the DSSF districts and other districts throughout the study period is necessary to support the claim that it was the DSSF program that produced higher high school test scores. In addition, we show that the greater increases occurred in categories of expenditures that can be plausibly linked to higher test score increases. So there are plausible mechanisms through which DSSF may have produced the discontinuity. And finally, we show that there was no interdistrict discontinuity in several other variables that might offer alternative explanations for the observed effect on high school test scores, which provides strong evidence that no hidden treatments occurred that might account for observed differences in test scores.

In the year before the DSSF pilot was initiated (2003–2004), overall per pupil expenditures in the 16 DSSF pilot districts were slightly higher than in all non-DSSF North Carolina districts as well as in the 16 North Carolina districts closest to the DSSF districts on the index of educational advantage used to select the DSSF pilot districts (for brevity, henceforward “closest” districts). Overall per pupil expenditures were $7,304 in DSSF pilot districts, $6,990 in all non-DSSF districts, and $7,266 in the 16 closest districts, as shown in Figure 1. Thus, greater percentage increases in overall per pupil expenditures in DSSF pilot districts than in other districts would also reflect greater increases in actual dollars spent. (Also note that DSSF districts are poor and largely rural.)

The percentage increase in overall per pupil expenditures from 2003–2004 to 2004–2005 in DSSF pilot districts (9.13%) was 4.85 percentage points higher than in non-DSSF districts (4.28%) and 4.41 points higher than in the 16 closest districts (4.72%). So there was a clear difference in the rate of increase in overall per pupil expenditures between DSSF districts and the other two sets of districts. This difference parallels the discontinuity in the relationship between the DSSF index of educational advantage and adjusted high school student test scores in DSSF versus other districts. The 4.85 percentage point difference in the first-year increase in overall per pupil expenditure between DSSF and non-DSSF districts also exceeds DSSF expenditures as a percentage of the pilot districts’ total budgets for 2004–2005 (3%)—evidence that the DSSF pilot districts did not use the new funds to supplant local funds.

The discontinuity in spending increases held up over the entire period of the pilot from the year prior to the DSSF pilot beginning (2003–2004) through the second year of the pilot (2005–2006). Between 2003–2004 and 2005–2006, overall per pupil expenditures increased by 13.81% in DSSF pilot districts but by only 8.28% in non-DSSF districts—a difference of 5.53 percentage points. During the same period, spending increased in the 16 closest districts by 8.57%, which represents a discontinuity in the rate of increased expenditures of 5.24 percentage points at the cutoff.

Moreover, the discrepancies in the rates of increase occurred in several categories of expenditures that might plausibly have contributed to increased high school test scores. These include regular classroom instruction, professional development, and instructional support. There were also differences in rates of increase in other categories, but these three expenditure categories can be plausibly linked to increased achievement.

**Regular Classroom Instruction**

Regular classroom instruction includes teachers’ salaries and bonuses as well as instructional
materials for non-special-education students. This category accounted for almost half of DSSF expenditures (46% in 2004–2005). In North Carolina, it is important to point out that districts are allowed to supplement teachers’ salaries above the amounts established on the state salary schedule. In the first year of the DSSF pilot (2004–2005), regular classroom per pupil expenditures increased by 10.16% in pilot districts but by only 4.01% in all non-DSSF districts and by only 4.15% in the 16 closest districts, differences of 6.15 and 6.01 percentage points, respectively, as shown in Figure 2. Over the entire study period, 2003–2004 to 2005–2006, the difference between the rate of increase in DSSF and all non-DSSF districts was 4.28 points, whereas that between DSSF and the closest districts was 3.27 points.

There are several ways in which increases in per pupil expenditures for regular classroom instruction may have contributed to improvements in student performance. They may have enabled DSSF districts to reduce teacher turnover rates, reducing the erosion of social relationships between teachers and students, shared sense of mission, and professional community in their schools. In fact, teacher turnover did decrease in the DSSF districts. Teacher turnover in the pilot districts fell from 28.5% to 25.6% (a 2.9 percentage point decrease), whereas the turnover rate in the other districts in the state slightly increased from 22.0% to 23.8%. The increases in instructional expenditures, including teacher bonuses, may also have improved teachers’ morale and motivation. In addition, they may have enabled principals to keep and hire more able, better motivated teachers.

**Professional Development**

Professional development expenditures have sometimes been defined in the literature to include both teacher training activities and instructional support such as media or technology expenditures (Killeen, Monk, & Pleck, 2002). In this study, we separate instruction of support (see below) from professional development, which includes expenditures for workshops or professional development activities, substitute teachers’ salaries and benefits, and the salaries and benefits for teachers mentoring new teachers. In the first year of the pilot, in DSSF districts the increase in per pupil expenditures for professional development was huge—44.33%. In non-DSSF districts, the increase in per pupil expenditures for professional development was 17.52%. In the 16 closest districts, the increase in per pupil expenditures for professional development was 23.23%. Over the entire study period from 2003–2004 to 2005–2006, the differences in the rate of increase in per pupil expenditures for professional development were narrower but still substantial, as shown in Figure 3. During the 2-year study period, DSSF pilot districts increased professional development expenditures by 50.84%, non-DSSF pilot districts by 21.90%, and the 16 closest districts by 21.07%. With strong state and federal accountability pressures in place, districts had an incentive to focus professional development funds on improving test scores. So it is plausible that DSSF districts’ advantage in increased PD expenditures produced enhancements to human capital, which in turn translated into better learning outcomes.

**Instructional Support**

Instructional support includes expenditures related to instructionally focused computer labs, library and media services, instructionally
related technical assistance for teachers, and salaries and benefits for instructionally related technology support personnel. In the first year of the pilot, DSSF districts increased instructional support expenditures by 30.79%, all non-DSSF districts by 5.91%, and the 16 closest districts by 11.13%. Over the entire 2003–2004 to 2005–2006 period, as shown in Figure 4, DSSF districts increased spending for instructional support by 34.43%, all non-DSSF districts by 11.96%, and the 16 closest district by 16.61%. The differentials in increases in spending for instructional support can be plausibly linked to improvements in the schools’ capacity to produce higher test scores in DSSF high schools.

Turning now to the possibility of hidden treatments, five initiatives or other policies, including state and federal accountability policies, could provide alternative explanations for any observed effects attributed to the DSSF pilot if these policies were disproportionately directed toward the pilot districts during the study period. As Imbens and Lemieux (2008) point out in their review of RD designs, “A first concern about RD designs is the possibility of other changes at the same cutoff value of the covariate” (p. 17). There were five policies or initiatives that took place in North Carolina during the study period that could have affected high school test scores in the pilot districts differently from the other districts if there was a sharp discontinuity in any of these “treatments” at or around the cutoff for DSSF funding: (a) state accountability pressures, (b) technical assistance team interventions by the NCDPI in low performing schools, (c) pressures from the court in the state’s school finance case known as the Leandro case, (d) No Child Left Behind (NCLB, 2002) accountability pressures, and (e) interventions by the Gates Foundation–funded New Schools Project. As we show below, none of the five appear to have affected the districts below the cutoff differently than those immediately above the cutoff.

**State Accountability Pressures**

North Carolina’s ABCs system of assessment-based school accountability applies to all schools throughout the state and thus would seem an unlikely cause of the discontinuity. But districts with more lower performing schools experience disproportionate pressures to improve, resulting in a differential impact of this apparently uniform statewide policy. North Carolina designates two categories of poorly performing schools: “low performing schools” and “priority schools.” Low performing schools...
have fewer than 50% of their students at grade level. In high schools, this means that less than 50% of the EOC tests taken resulted in passing or “proficient” scores. A school can be designated a “priority school” for one of two reasons: (a) less than 60% but more than 50% of tests taken are passed (b) or less than 50% are passed but, on average, students learned as much during a given year as they were expected to learn (“expected growth”). During the first year of the DSSF pilot (2004–2005), the DSSF pilot districts, the 16 closest districts, and all non-DSSF districts had less than 1% of their schools designated as low performing and priority schools. In 2005–2006, DSSF districts actually had lower percentages of low performing schools (2.2) than the 16 closest districts (3.7) and the same proportion as all non-DSSF districts (2.3). That year, the DSSF districts had a lower percentage of priority schools (7.1) than both the 16 closest districts (14.7) and all non-DSSF districts (10.6). There were some differences, but the direction of the differences does not support the view that the state’s accountability pressures were more likely to have been felt in the DSSF pilot districts and certainly no sharp discontinuity between DSSF and other districts occurred in their percentages of low performing and priority schools. Thus, differential effects of the state accountability system could not account for the observed discontinuity in the relationship between the index of educational advantage and adjusted high school test scores.

**NCDPI “Assistance Teams”**

During the years of the DSSF pilot, the NCDPI regularly dispatched “assistance teams” to low performing schools. If schools in DSSF districts had received NCDPI assistance at sharply higher rates from the 16 closest districts, this might account for the discontinuity in the high school test scores at the cutoff. But as we have just shown, there was no such sharp difference between DSSF and other districts in the percentage or number of low performing schools and thus there would have been no sharp difference with respect to NCDPI assistance teams.

**The Leandro School Finance Case**

In the spring of 2005 Superior Court Judge Howard Manning—the cognizant judge in North Carolina’s long-running *Leandro* school finance case—severely criticized high schools with performance composites (passing rates on the state EOC exams) consistently below 55% and threatened to close those that did not bring their performance up immediately. In response, the NCDPI sent special “turnaround teams” into these especially low performing high schools. If Manning had targeted substantially more high schools in DSSF districts than in the 16 closest districts, the additional pressure and support might have accounted for the apparent DSSF effect. But in fact Manning’s list included only 7 high schools in DSSF districts, compared to 12 high schools in the 16 closest districts.

**NCLB Accountability Pressures**

Districts with disproportionate numbers of schools failing to make adequate yearly progress (AYP) as required by the federal NCLB would also experience disproportionate pressures to improve. As Table 1 reflects, from 2003–2004 through 2006–2007, there were indeed major differences between the percentage of schools failing to make AYP in DSSF versus all non-DSSF districts. But the differences between DSSF districts and the 16 closest districts were generally quite small. NCLB may have increased the pressure on DSSF districts to use the new funds to improve student test scores, but it seems unlikely that differential NCLB pressures themselves account for the sharp discontinuity in the relationship between the index of educational advantage and adjusted high school test scores.

**New Schools Project Schools**

The North Carolina New Schools Project (NCNSP) is a private, nonprofit high school reform organization initiated in 2003 by Governor Mike Easley with support from the Gates Foundation. The New Schools Project has now supported the development of more than 100
TABLE 1  
Schools Failing to Make Adequate Yearly Progress Comparisons: DSSF Pilot Districts, 16 Closest Districts, and All Non-DSSF Districts

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DSSF districts</td>
<td>38</td>
<td>58</td>
<td>68</td>
</tr>
<tr>
<td>16 closest districts</td>
<td>37</td>
<td>53</td>
<td>65</td>
</tr>
<tr>
<td>All non-DSSF districts</td>
<td>28</td>
<td>41</td>
<td>54</td>
</tr>
</tbody>
</table>

Note. DSSF = Disadvantaged Student Supplemental Fund.

innovative high schools statewide, but according to the NCNSP website, the first of these was created in 2005–2006, too late to have affected outcomes during the first year of the DSSF pilot (2004–2005). Furthermore, in 2005–2006 the NCNSP actually initiated more innovative high schools in the 16 closest districts (5) than in DSSF districts (1). So the New Schools Project interventions cannot have brought about the discontinuity in high school test scores at the cutoff.

To the best of our knowledge, there were neither other policy developments in the North Carolina nor any differences in secular trends that might account for the observed discontinuity in the relationship between the index of educational advantage and adjusted high school test scores. In summary, then, we have documented that (a) the DSSF pilot resulted in a real treatment (a differential increase in total expenditures between DSSF and other districts, including the 16 districts closest to the cutoff), (b) there are plausible mechanisms through which the treatment could have brought about the observed discontinuity, and (c) to a reasonable level of certainty, there were no hidden treatments.

4. Study Design, Sample, Measures, and Estimation Strategy

For evaluations assessing the impact of education policies, a primary goal is that the design, sample, measures, and analysis combine to produce an unbiased estimate the magnitude of the policy’s effect on important measures of expected outcomes for a group that is relevant to future policy decisions. In this study, we use a RD design and mixed models to identify and estimate the effects of a targeted funding program on high school student achievement in a large and diverse state. Below we unpack this summary of the research design, sample, and data used for our analyses.

A. Research Design

The primary research questions for this study are the following:

1. Can additional funds targeted to districts with the fewest educational advantages increase high school student achievement?
2. Can additional funds targeted to districts with the fewest educational advantages increase the achievement of academically disadvantaged high school students?

Addressing these questions requires that the impact of the effects of the intervention, in this case targeted funding such as that provided by the North Carolina DSSF pilot, be estimated without bias, using Rubin’s (1974, 2005) terminology. In this context, this means that the estimate of the program’s effect not be biased by unequal distribution of disturbing yet uncontrolled factors between the group that participates in the intervention and the group that does not participate. From a theoretical standpoint, a design that includes random assignment to a group that will participate in the intervention or a group that will not participate equivalently distributes the disturbing factors to the two groups. Correctly estimated, the comparison of the postintervention outcomes for the randomly assigned participants and nonparticipants can yield unbiased estimates of the effects (Holland, 1986) for the population from which they were...
drawn. Unquestionably, other sources of bias can arise after random assignment or because of choices of the study sample or measures, but random assignment designs have the capacity to produce unbiased estimates of policy effects.

As a precursor to a random assignment study or in situations when a random assignment study is proven to be infeasible or unethical, other designs and various types of statistical controls have been proposed to estimate causal effects. These designs require more assumptions, both testable and untestable, than random assignment studies for the estimation of causal effects. But matching designs, including propensity score matching, and RD designs have shown promise for reducing or eliminating bias more than regression-based approaches alone (Cook et al., 2008; Glazerman et al., 2003). Specifically, RD designs can be employed to effectively eliminate selection bias—the most pernicious form of bias that can be addressed through design—from the estimates of effects (Imbens & Lemieux, 2008; Mark & Shotland, 1987; Shadish et al., 2001; Trochim, Cappelleri, & Reichardt, 1991; van der Klaauw, 2002). As Schochet (2008) indicates in a recent publication,

Under well-designed RD designs, the treatment assignment rule is fully observed and can be modeled to yield unbiased impact estimates. . . . An impact occurs if there is a “discontinuity” in the two regression lines at the cutoff score. Because the selection rule is fully known under the RD design, selection bias issues tend to be less problematic under the RD design than under other non-experimental designs. The literature suggests that the RD design might be a suitable alternative to a random assignment (RA) design when an experiment is not feasible (Cook, 2008). (p. 1)

Cook et al. (2008) provide evidence that the differences in estimates from RD and random assignment studies are negligible. RD design eliminates selection bias by including the appropriate functional form of the assignment variable, which can be accomplished by adding higher order terms of the assignment variable in the estimating equation (Imbens & Lemieux, 2008, p. 10; Shadish et al., 2001; Trochim et al., 1991; van der Klaauw, 2002). Including variables in the estimation equation that perfectly model the relationship between selection and the outcome variable, including nonlinear relationships, breaks up any possible correlation between the variable indicating participation in the intervention and the disturbance term, thereby removing selection bias from the estimate of the treatment effect.

In this study, an index of educational advantage was used to create a sharp discontinuity between the districts that received the intervention and those that did not. In the following equations, we show the theory that underlies RD elimination of selection bias. A typical value added equation to estimate the treatment effect, \( \beta_i \), is shown in Equation 1,

\[
Y_{it} = \beta_0 + \beta_1 \text{DSSF}_j + \gamma_n Y_{i0-n} + \gamma_y X_y + \mu_i
\]  

(1)

where \( i \) indexes individual students, \( j \) indexes districts, \( t \) indexes time, \( Y_{it} \) represents a student test score, \( \text{DSSF}_j \) is the indicator variable for assignment to the intervention, \( Y_{i0-n} \) is a vector of prior test scores, \( X_y \) is a vector of covariates, and \( \mu_i \) is the disturbance term. The estimate of the treatment effect in Equation 1, \( \beta_i \), will be biased if the selection into the program is correlated with omitted variables that also influence the test scores, \( Y_{it} \). In Equation 2 we add terms for a polynomial function of \( I_j \), the index of educational advantage.

\[
Y_{it} = \beta_0 + \beta_1 \text{DSSF}_j + f(I_j) + \gamma_n Y_{i0-n} + \gamma_y X_y + \mu_i
\]  

(2)

Basically, the \( I_j \) terms will “soak up” the selection bias that may be attributable to omitted variables, so that although their coefficients may be biased, the coefficient indexing the intervention effect will be correspondingly purged of selection bias. The elimination of selection bias requires that the assignment to treatment is perfectly measured, which in the case of the DSSF program was true. The covariates are included to eliminate small sample biases and increase precision without altering the identification strategy, as Imbens and Lemieux (2008, p. 11) state.

As Schochet (2008) indicates, a well-designed RD study effectively eliminates selection bias, but researchers must take care to assess other study limitations. Just as estimates of program effects from random assignment studies can be biased from events that occur after randomization such as differential attrition,
crossovers from treatment to control and vice versa, and researcher expectancies, RD is not immune to threats to validity (Reichardt & Henry, in press). A primary validity concern for RD designs is the generalizability of the estimate of the effect of the intervention. RD designs estimate the effect of the intervention at the cutoff, which can be different from the average treatment effects obtained in random assignment studies. This could present a threat to external validity (Shadish et al., 2001) and could reduce confidence in extrapolating the effects to other districts. In this study, we carefully limited the generalizability of the effects to the most educationally disadvantaged districts and did not extrapolate the effects of increasing funding to all school districts. In addition, we estimated the impact of DSSF with alternative samples to assess possible limits on the extrapolation of the effects. Finally, we carefully assessed the potential for hidden treatments or abrupt changes in observed covariates that may have brought about differences in student achievement at the cutoff. Following the recommendation of Imbens and Lemieux (2008) to graphically examine the changes in covariates at the cutoff, we did so and did not find confounding covariates that were discontinuous at or around the cutoff.

B. Sample

For this study, we assembled data from all 337 regular public high schools that operated in North Carolina in 2004–2005 and 2005–2006. In these 2 years, a total of 414 public schools offered educational programs in grades 9 to 12. We defined schools as regular public high schools if they normally offered courses required for high school graduation, were not designated as vocational, technical, or career prep high schools, and were not listed as alternative high schools. For these 337 high schools we include scores on EOC exams that are given just prior to completion of eight courses: algebra 1, algebra 2, geometry, biology, chemistry, physical science, physics, and English 1. Over the course of the 2 study years, approximately 230,000 students took these tests. All of the students were included in our data, but in some of the analyses a small percentage of cases were omitted because of missing data.

C. Measures

Five types of variables are included in this study: student EOC scores, student control variables, classroom- and school-level control variables, the index of educational advantage, and an indicator of assignment to the pilot program.

Student EOC scores. As previously mentioned, high school students take EOC exams to assess the extent to which they have acquired the knowledge and skills expected by the state for the particular subject. For this study, these tests were developed for the state using methods that meet both validity and reliability criteria for high stakes tests. The North Carolina EOC tests are used to sample a student’s knowledge of subject-related concepts as specified in the North Carolina Standard Course of Study and to provide a global estimate of the student’s mastery of the material in a particular content area. The purpose of the EOC tests is threefold: to ensure that all high school graduates possess the minimum skills and knowledge thought necessary to function as a member of society, to provide a means of identifying strengths and weaknesses in the educational process for the purpose of instructional improvement, and to provide information for the educational accountability systems at the state, local, and school levels. The North Carolina EOC tests were initiated in response to the North Carolina Elementary and Secondary Reform Act of 1984. In the years covered by our data, students enrolled in algebra 1, algebra 2, biology, chemistry, English 1, geometry, physics, and physical science courses were required to take the North Carolina EOC tests. In addition, the combined passing rate for these tests, which is called the performance composite, is the principal means by which the performance of high schools is assessed by the state’s accountability system. The test development process for these North Carolina curriculum-based, criterion-referenced tests is rigorous, includes input from teachers and subject matter specialists, and adheres to standards for test development set by the U.S. Department of Education for implementation of NCLB (2002).

The scores for the multiple-choice EOC tests are reported as scale scores, percentiles, and achievement levels. The scale scores measure
Individual-student-level controls. For the purpose of this study, we assembled an extensive set of individual student variables that were collected for high school students in the state. Prior achievement was measured by each student’s eighth grade reading and mathematics tests. Although there is a continuing exploration of the preferred dependent variable for EPF studies, gain scores and value added equations in which the prior achievement is included as a control are the most discussed options. Todd and Wolpin (2003) show that, holding all other variables constant, the former is a special case of the latter in which the coefficient on the prior achievement score is constrained to equal 1. In this case, because we have EOC tests that assess different subjects and we have two scores for prior achievement, we used the value added specification with both measures of prior achievement included as controls.

We include peer ability within the classroom and student absences as additional controls. Peer ability is measured as the average eighth grade test score (reading and mathematics) of the other students in the individuals’ classroom for each of their courses for which an EOC score is included. In addition, we added dichotomous variables for race/ethnicity: African American, Hispanic, Asian, Native American or American Indian, and Multiracial (White is the reference category). As an indicator of family resources, we include separate dichotomous variables for eligibility for free lunches and eligibility for reduced price lunches as well as an indicator for students who do not have free or reduced by lunch eligibility or ineligibility recorded (not eligible for free or reduced and not missing is the reference category). In addition, we included indicator variables for exceptional students by category, including gifted and six categories of disabilities: high incidence, cognitive, behavioral, sensory, physical, and severe disabilities. We also include two variables that indicate whether the student currently or previously received services for LEP. Parental education is included for four categories of education: less than high school, high school only, college graduate, and missing (some college is the reference group). The missing parental education variable was included to avoid biasing the sample by omitting students whose parental education was not included in the administrative data. We include two other categories of student variables, one measuring age and the other grade level. We include indicators of underage or overage for the student’s grade level to index students who may have been promoted two grades at once or started school early and those who have been retained in grade at least once, respectively. Finally, we include whether the student was a sophomore, junior, or senior at the time he or she took the EOC exam, with freshman being excluded as the reference group.

Additional classroom and school variables. School-level variables include a group of compositional variables that are mainly used to control for school and classroom contextual effects. Seven dichotomous variables at the individual level were aggregated to create school percentages by race/ethnicity and as well as by free lunch eligibility and reduced price lunch eligibility. In addition, we include school size as measured by average daily membership. At the classroom level, we include indicator variables for whether the class was remedial or advanced. Finally, we include a measure of the dispersion of the ability level of the pupils in each EOC class, computed as the standard deviation of the students’ average eighth grade test score (reading and mathematics).

Index of educational disadvantage. As noted in the description of the DSSF pilot program, the index of educational advantage used for assignment to the pilot included four (equally weighted) measures of educational advantage: teacher stability, experienced teachers, children not living...
in poverty, and students meeting state proficiency standards. We include this measure centered at the cut-off as well as higher order terms in the estimating equations to produce consistent (unbiased) estimates of the program’s impact.

**Program indicator.** The variable of interest for the evaluation of program effects is the variable that indicates participation in the DSSF program. Because the assignment to the intervention was based strictly on the value on the education advantage index, the coefficient on the variable of interest assuming a proper specification of the analytical model is an unbiased estimate of the program’s impact on high school student achievement. The analytical model is discussed in the next section.

**D. Analytical Model**

The analysis estimating the effects of targeted funding on student achievement was conducted using a multilevel model to allow for nesting of students within classrooms, schools, and districts. The models estimated for this study are shown in Equation 3:

\[
Y_{it} = \beta_0 + \beta_1 DSSF_i + f(I_{it}) + \gamma_s Y_{it-1} + \gamma_x X_i + \gamma_Z Z_i + \gamma_W W_i + \mu_i + \epsilon_i + \theta_i
\]

In Equation 3, \(Y_{it}\) is the score for the student \(i\) on EOC test \(t\) in school \(s\); \(\beta_1\) is the coefficient on the intervention indicator variable. Vectors of lagged test scores, student characteristics, classroom controls, and school controls are symbolized by \(Y_{it-1}, X, Z,\) and \(W,\) respectively. The stochastic portion of the model is symbolized by \(\mu, \epsilon,\) and \(\theta,\) which represent the disturbance terms at the individual, classroom, and school levels. For inference, following Imbens and Lemieux (2008, p. 19), robust standard errors that accounted for nesting within schools and classrooms were used. Descriptive statistics for each of the variables included in the model are included in the appendix. The final model including all students contained more than 813,000 student EOC scores and 56,414 classroom observations over the 2-year period. The amounts of variance of the full sample for the null model at the school, classroom, and student levels are 20.0%, 35.5%, and 44.4%, respectively, indicating that substantial explanatory power is added by using mixed models rather than aggregated data.

**5. Study Findings**

In this section, we first provide an overview of our estimates of the effects of DSSF on high school student achievement. Then we present details of our findings that the DSSF pilot did improve high school student test scores on average and for academically disadvantaged students in pilot districts.

**A. Overview of the Findings**

High school students in the DSSF districts score approximately 0.13 (\(p < .001\)) of a standard deviation higher than the students in non-DSSF districts, controlling for student, classroom, and school characteristics, including selection into the pilot. This corresponds to an overall difference of about 1.2 points in average EOC scores, which we graphically illustrate in Figure 5. On the horizontal axis, we plotted the educational advantage index for districts in the state; and on the vertical axis, we plotted the adjusted average EOC score. To the left of the vertical line, the graph displays the DSSF districts and to the right the non-DSSF districts, which have higher scores on the index of educational advantage. The solid line indicates the expected average EOC score, adjusted for district characteristics, which would have been expected in the absence of DSSF. The dotted 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 discontinuity between the two lines for the pilot districts indicates the size of the effect because of DSSF. The break in the solid line for the other districts on one hand and the DSSF districts on the other indicates the program impact at the point that was chosen as the cutoff for receiving the supplementary funding. After a very close inspection of alternative causes as reported earlier in the article and several checks on robustness, we find no other explanation for the sharp difference in high school achievement that occurs at precisely the cutoff point.

Our second research question focused attention on academically disadvantaged students.
Adjustment to Scores

Adjusted Scores DSSF Districts

Adjusted Scores Non-DSSF Districts

Poly (Fitted Values - Excluding DSSF Effect)

Poly (Fitted Values - Including DSSF Effect)

FIGURE 5. District average test score by educational advantage index. Note. DSSF = Disadvantaged Student Supplemental Fund; LEA = Local Education Agency.

TABLE 2A
Estimates of DSSF Program Effects for All Students and Academically Disadvantaged Students

<table>
<thead>
<tr>
<th>1. Variables</th>
<th>2. All students</th>
<th></th>
<th>3. Academically disadvantaged students</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td></td>
<td>SE</td>
</tr>
<tr>
<td>DSSF indicator</td>
<td>0.1332</td>
<td>(0.0344)**</td>
<td>0.0984</td>
<td>(0.0363)**</td>
</tr>
<tr>
<td>Educational index</td>
<td>0.3904</td>
<td>(0.1149)**</td>
<td>0.2303</td>
<td>(0.1297)</td>
</tr>
<tr>
<td>Educational index squared</td>
<td>0.5767</td>
<td>(0.2542)*</td>
<td>0.0999</td>
<td>(0.3501)</td>
</tr>
<tr>
<td>Educational index cubed</td>
<td>-2.3494</td>
<td>(0.7268)**</td>
<td>-0.4011</td>
<td>(0.9474)</td>
</tr>
<tr>
<td>School-level covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage Asian</td>
<td>-0.0015</td>
<td>(0.0025)</td>
<td>-0.0043</td>
<td>(0.0026)</td>
</tr>
<tr>
<td>Percentage Black</td>
<td>0.0004</td>
<td>(0.0004)</td>
<td>0.0008</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Percentage Hispanic</td>
<td>0.0019</td>
<td>(0.0016)</td>
<td>0.0021</td>
<td>(0.0015)</td>
</tr>
<tr>
<td>Percentage Multiracial</td>
<td>0.0075</td>
<td>(0.0043)</td>
<td>-0.0033</td>
<td>(0.0051)</td>
</tr>
<tr>
<td>Percentage American Indian</td>
<td>-0.0009</td>
<td>(0.0008)</td>
<td>-0.0003</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Percentage free lunch</td>
<td>-0.0010</td>
<td>(0.0007)</td>
<td>-0.0014</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>Percentage reduced lunch</td>
<td>0.0011</td>
<td>(0.0014)</td>
<td>0.0015</td>
<td>(0.0015)</td>
</tr>
<tr>
<td>School size</td>
<td>0.0000</td>
<td>(0)</td>
<td>0.0000</td>
<td>(0)</td>
</tr>
<tr>
<td>Indicator for 2005–2006</td>
<td>-0.0376</td>
<td>(0.01)**</td>
<td>-0.0227</td>
<td>(0.0108)*</td>
</tr>
<tr>
<td>Classroom-level covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced curriculum</td>
<td>0.1439</td>
<td>(0.0091)**</td>
<td>0.1958</td>
<td>(0.0214)**</td>
</tr>
<tr>
<td>Remedial curriculum</td>
<td>0.0410</td>
<td>(0.0295)</td>
<td>0.0519</td>
<td>(0.0364)</td>
</tr>
<tr>
<td>Classroom ability dispersion</td>
<td>0.0274</td>
<td>(0.0132)*</td>
<td>-0.0604</td>
<td>(0.0209)**</td>
</tr>
</tbody>
</table>

Note. DSSF = Disadvantaged Student Supplemental Fund.
*p < 0.05. **p < 0.01.
TABLE 2B  
*Estimates of the Impact of the DSSF Pilot Program: Individual-Student-Level Covariates*

<table>
<thead>
<tr>
<th>1. Variables</th>
<th>2. All students</th>
<th></th>
<th>3. Academically disadvantaged students</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>Coeff.</td>
<td>SE</td>
</tr>
<tr>
<td>Individual-level controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8th grade reading score</td>
<td>0.2538</td>
<td>(0.0017)**</td>
<td>0.2379</td>
<td>(0.0046)**</td>
</tr>
<tr>
<td>8th grade mathematics score</td>
<td>0.4422</td>
<td>(0.0026)**</td>
<td>0.3040</td>
<td>(0.0057)**</td>
</tr>
<tr>
<td>Absences</td>
<td>-0.0086</td>
<td>(0.0002)**</td>
<td>-0.0065</td>
<td>(0.0003)**</td>
</tr>
<tr>
<td>Classroom peer ability</td>
<td>0.1751</td>
<td>(0.0072)**</td>
<td>0.1901</td>
<td>(0.0091)**</td>
</tr>
<tr>
<td>Male (= 1)</td>
<td>0.0258</td>
<td>(0.0021)**</td>
<td>-0.0039</td>
<td>(0.0053)</td>
</tr>
<tr>
<td>Asian (= 1)</td>
<td>0.0454</td>
<td>(0.0076)**</td>
<td>0.0715</td>
<td>(0.0257)**</td>
</tr>
<tr>
<td>Black (= 1)</td>
<td>-0.1002</td>
<td>(0.003)**</td>
<td>-0.1133</td>
<td>(0.0063)**</td>
</tr>
<tr>
<td>Hispanic (= 1)</td>
<td>-0.0095</td>
<td>(0.0059)</td>
<td>-0.0059</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>Multiracial (= 1)</td>
<td>-0.0167</td>
<td>(0.0063)**</td>
<td>-0.0144</td>
<td>(0.0198)</td>
</tr>
<tr>
<td>American Indian (= 1)</td>
<td>-0.0636</td>
<td>(0.0098)**</td>
<td>-0.1062</td>
<td>(0.0161)**</td>
</tr>
<tr>
<td>Eligible for free lunch</td>
<td>0.0149</td>
<td>(0.0025)**</td>
<td>0.0037</td>
<td>(0.0055)</td>
</tr>
<tr>
<td>Eligible for reduced-price lunch</td>
<td>0.0159</td>
<td>(0.0036)**</td>
<td>0.0159</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>Eligibility free or reduced missing</td>
<td>0.0220</td>
<td>(0.0203)</td>
<td>-0.0013</td>
<td>(0.0429)</td>
</tr>
<tr>
<td>Gifted</td>
<td>0.1141</td>
<td>(0.0042)**</td>
<td>0.1013</td>
<td>(0.0782)</td>
</tr>
<tr>
<td>High incidence</td>
<td>-0.0431</td>
<td>(0.0049)**</td>
<td>-0.0728</td>
<td>(0.0076)**</td>
</tr>
<tr>
<td>Cognitive</td>
<td>-0.1212</td>
<td>(0.0157)**</td>
<td>-0.1993</td>
<td>(0.0153)**</td>
</tr>
<tr>
<td>Behavioral</td>
<td>-0.0144</td>
<td>(0.0179)</td>
<td>-0.0238</td>
<td>(0.0213)</td>
</tr>
<tr>
<td>Sensory</td>
<td>-0.0152</td>
<td>(0.0212)</td>
<td>-0.0553</td>
<td>(0.0375)</td>
</tr>
<tr>
<td>Physical</td>
<td>-0.0858</td>
<td>(0.0102)**</td>
<td>-0.1152</td>
<td>(0.0144)**</td>
</tr>
<tr>
<td>Severe</td>
<td>0.0339</td>
<td>(0.0114)**</td>
<td>-0.0107</td>
<td>(0.0217)</td>
</tr>
<tr>
<td>Previous LEP</td>
<td>-0.0112</td>
<td>(0.013)</td>
<td>-0.0305</td>
<td>(0.0193)</td>
</tr>
<tr>
<td>Current LEP</td>
<td>0.0408</td>
<td>(0.0167)*</td>
<td>0.0952</td>
<td>(0.0424)*</td>
</tr>
<tr>
<td>Parent ed. &lt; high school</td>
<td>-0.0208</td>
<td>(0.0041)**</td>
<td>-0.0482</td>
<td>(0.0088)**</td>
</tr>
<tr>
<td>Parent ed. = high school</td>
<td>-0.0216</td>
<td>(0.0022)**</td>
<td>-0.0426</td>
<td>(0.0054)**</td>
</tr>
<tr>
<td>Parent ed. = college graduate</td>
<td>0.0129</td>
<td>(0.0025)**</td>
<td>-0.0146</td>
<td>(0.0063)*</td>
</tr>
<tr>
<td>Parent ed. missing</td>
<td>-0.0870</td>
<td>(0.0158)**</td>
<td>-0.1256</td>
<td>(0.027)*</td>
</tr>
<tr>
<td>Underage</td>
<td>0.1048</td>
<td>(0.007)**</td>
<td>0.1300</td>
<td>(0.027)*</td>
</tr>
<tr>
<td>Overage</td>
<td>-0.0803</td>
<td>(0.0027)**</td>
<td>-0.0944</td>
<td>(0.0051)**</td>
</tr>
<tr>
<td>Grade 10</td>
<td>0.0007</td>
<td>(0.005)</td>
<td>0.0398</td>
<td>(0.0098)*</td>
</tr>
<tr>
<td>Grade 11</td>
<td>0.0116</td>
<td>(0.0069)</td>
<td>0.0888</td>
<td>(0.0109)**</td>
</tr>
<tr>
<td>Grade 12</td>
<td>0.0825</td>
<td>(0.0084)**</td>
<td>0.1636</td>
<td>(0.0144)**</td>
</tr>
<tr>
<td>English 1</td>
<td>0.0379</td>
<td>(0.0095)**</td>
<td>-0.0363</td>
<td>(0.0125)**</td>
</tr>
<tr>
<td>Algebra 2</td>
<td>-0.3299</td>
<td>(0.011)**</td>
<td>-0.2933</td>
<td>(0.0159)**</td>
</tr>
<tr>
<td>Physical science</td>
<td>0.1535</td>
<td>(0.0119)**</td>
<td>0.0148</td>
<td>(0.0134)</td>
</tr>
<tr>
<td>Biology</td>
<td>-0.0639</td>
<td>(0.0096)**</td>
<td>-0.1240</td>
<td>(0.0118)**</td>
</tr>
<tr>
<td>Chemistry</td>
<td>-0.5876</td>
<td>(0.0152)**</td>
<td>-0.6299</td>
<td>(0.0245)**</td>
</tr>
<tr>
<td>Physics</td>
<td>-1.0144</td>
<td>(0.024)**</td>
<td>-1.0531</td>
<td>(0.0865)**</td>
</tr>
<tr>
<td>Geometry</td>
<td>-0.2745</td>
<td>(0.0091)**</td>
<td>-0.3684</td>
<td>(0.0132)**</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.0379</td>
<td>(0.0095)**</td>
<td>-0.0363</td>
<td>(0.0125)**</td>
</tr>
</tbody>
</table>

*Note. DSSF = Disadvantaged Student Supplemental Fund; LEP = limited English proficiency.  
*p < 0.05. **p < 0.01.*

Academically disadvantaged high school students are students now enrolled in the state’s 337 regular high schools who entered high school not proficient in either math or reading. These students were the explicit focus of the pilot funding program. Relying on the same specification as above for the disadvantaged student sample, the DSSF effect was estimated.
to be 0.10 standard deviation units, a slightly smaller and statistically significant effect.

Column 2 in Table 2a displays the results from the specification of the model using the DSSF indicator variable as the variable of interest for all students, and column 3 displays the estimate for the academically disadvantaged students.

For both equations, higher order terms of the index of educational disadvantage were tested for significance. The highest order term that was significant was the third degree polynomial (cubed term), and the second degree term was sometimes significant. Therefore, we included the linear, squared, and cubed terms in the estimation equation, including in the robustness analysis that follows, even though the squared and cubed terms were never significant in the limited sample tests of robustness. Thus, it appears that the higher order terms corrected for nonlinearity in the relationship between the assignment variable and average EOC test scores in the full sample analysis that compared 16 districts to all other districts in the state.

After the effects were estimated using all the districts in the state, we examined the robustness of the original specification and measures of the treatment. Generally with RD studies, the effect estimates are reestimated with different bandwidths or bins containing percentages of cases on each side of the cutoff point on the assignment variable. Unfortunately, we were limited in our ability to test many different bandwidths because so few pilot districts fell below the cutoff. For the first check on robustness, we limited the comparison sample of districts to the 16 districts closest to 16 DSSF pilot districts on the assignment variable, the index of educational advantage. If the regression is sensitive to cases that are far away from the cutoff (the most educationally advantaged districts) to produce the difference and the inclusion of higher order terms in the estimating equation did not fully adjust for this, limiting the regression to 16 above and 16 below the lines should expose the problem. The effect estimated for the 32 districts using the same specification as the original estimate was 0.12 ($p < .05$). The efficiency of the estimate was reduced, as would be expected given the reduction in the sample size, but the estimate was very similar in size as the estimate on the full sample. We also estimated the effect reducing the sample from 32 districts to 16, 8 above and 8 below the cutoff. The effect with this restricted sample was no longer reliable; the standard error increased from 0.034 with the full sample to 0.061 with 16 districts, and the magnitude of the coefficient was reduced. In addition, we included an interaction term—the treatment indicator with the index variable—and reran the model with the interaction term along with the original specification, and the coefficient on the interaction term was not significant.

In addition to these checks on the stability of the effect estimates for different samples of districts, we assessed the temporal stability of the estimates. The estimating equation included a term for the second year of the pilot. Controlling for other variables, in the second year of the pilot, 2005–2006, average EOC test scores were a little lower than in the previous year. To check the stability of the estimates over the 2 years, we interacted the indicator variable for the second year of the pilot, 2007, with the indicator for DSSF districts. The interaction was not statistically significant, indicating that the effect was not temporally heterogeneous.

In another check on robustness, we used an alternative measure of the variable of interest to test the effect of the magnitude of additional DSSF funding on high school student performance in pilot districts. For this check on robustness, we replaced the dichotomous indicator of treatment with a continuous measure, the level of DSSF expenditures on a per-pupil basis at each high school. Recall that the amounts of per pupil funding were based on $250 per pupil at the district level, but because of the flexibility in how the funds were used, the actual amount for any high school could be more or less than the district per pupil allocation. Testing an alternative variable of interest could add to our confidence in the findings if the results are both positive and of a similar magnitude. The alternative variable would indicate that the amount of DSSF funds spent in the high schools relates to higher achievement. We included the third degree polynomial functions of assignment variable used to select the pilot districts and other controls exactly as was done in the original estimation. The results show that $100 in DSSF funds produced an increase of about 0.044 of a standard deviation in average scores. The effect
of the DSSF $250 was about 0.11 of a standard deviation, which approximates the 0.133 standard deviation effect from the DSSF indicator variable noted above. This helps to corroborate the first analysis and increases our confidence that the estimate of the program effect can be attributed to DSSF funding.

**B. Covariates: Individual Student Level**

Because achievement on the EOC tests also depends on student, classroom, and school characteristics, we included covariates to adjust for unequal distribution of these characteristics between the pilot districts and the other districts in the state and to increase precision (Imbens & Lemieux, 2008). At the individual student level, a variety of student characteristics are associated with a statistically significant effect on students’ average high school achievement scores, as shown in Table 2b. Among individual characteristics, student performance on eighth grade reading and mathematics assessments had the strongest associations with high school achievement. Scoring one standard deviation higher on eighth grade End of Grade (EOG) reading is associated with an increase of about 0.25 standard deviations on the high school EOC exams, holding other variables constant. An increase of one standard deviation on a student’s eighth grade EOG math score is associated with an increase in the student’s EOC test score of about 0.44 standard deviations for students similar in all other respects. The number of days a student is marked absent from school has a negative relationship with EOC test scores. An additional 10 absences lowers student performance by about 0.09 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 his or her high school classes (measured in eighth 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, few students have peers who vary a great deal from the average score, which makes the average increment closer to 0.03 standard deviations or about three tenths of a point on average.

On average, male students score slightly higher on high school EOC exams than comparable female students, by about 0.026 standard deviations. Holding constant other factors, when compared to White students, Asian students perform slightly better (0.04 SD), Black, Multiracial, and American Indian students perform slightly worse (−0.10, −0.02, and −0.06 SDs, respectively), and Hispanic students perform no differently from a statistical point of view.

High school test scores also vary by family income. Perhaps surprisingly, compared to students not eligible for federal lunch subsidies, students eligible for free lunch, students eligible for reduced price lunch, and students whose free or reduced lunch eligibility data were missing all performed slightly better (about 0.015, 0.016, and 0.022 SDs, respectively) when controlling for other variables in the model, including prior test performance. 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, those with high-incidence disabilities, and those with physical disabilities scored lower on average than comparable regular instruction students (about 0.012, 0.04, and 0.09 SDs, respectively). Students who have a behavioral or sensory disability performed as well as students with no coded disabilities. Finally, students with a severe disability actually performed slightly better on average than comparable regular instruction students. This finding is likely attributable to including only higher performing students with this type of disability in the EOC testing.

Students receiving 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 association may seem contrary to expectations, but keep in mind that a student’s eighth 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 could explain this difference in achievement.

The level of education attained by a student’s parents was also significantly related to the student’s high school student achievement. 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 SDs on average for both groups). Students whose parents were college graduates performed, on average, slightly better than those whose parents attended some college (0.013 SDs). Students whose parental education level was missing in the data scored about 0.09 standard deviations lower than students whose parents attended some college.

The coefficients for the underage and underage variables indicate the difference in EOC test achievement between similar students who differ on only the underage or underage variable. Underage students, those whose birthdates 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. Overage students, those whose birthdates would place them in a higher grade level given their age (they may have started school late or been previously retained), scored about 0.08 standard deviations lower compared to similar students of the appropriate age for their grade level.

C. Covariates: Classroom Level

Enrollment in a class with an advanced curriculum—honors or AP—has a positive association with EOC test scores, holding prior achievement and other control variables constant. 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. On average, high school students enrolled in a class with students who are one standard deviation more heterogeneous with respect to their prior achievement can be expected to score about 0.03 standard deviations higher on achievement tests than comparable students in a classroom where every student had the same score on their eighth grade tests (SD of zero).

D. Covariates: School Level

Net of the other variables in the model, the ethnic composition of a high school’s student population was not associated with the average EOC exam scores in the school, nor was the percentage of students receiving free or reduced price lunch. Similarly, controlling for all other variables in the model, the size of the high school was not associated with average EOC scores in the school.

E. Summary of the Effect Estimates

High school students in the 16 DSSF pilot districts scored significantly higher on their EOC achievement tests than would have been expected in the absence of the program. On average, high school students in the DSSF districts scored about 1.2 points higher on the high school exams than they would have scored without the program. The comparable boost for academically disadvantaged students was smaller, approximately 0.10 standard deviation units, or a little more than 0.9 points. Both effects were statistically significant, and the sizes of the effects held up in several checks on their robustness.

6. Study Conclusions and Policy Implications

Our findings show that the supplemental funding targeted to disadvantaged districts exerted its effects broadly rather than in a highly targeted manner—we found improved performance for the all student sample, not just for the disadvantaged student sample. The supplemental funds were effective at reducing high school performance gaps between disadvantaged districts and other districts. 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 was about 3.7 points. Without DSSF, the expected difference in scores would have been about 4.9 points.

The DSSF was also effective in reducing the gap between academically disadvantaged students in the pilot districts and similar students in other districts. For just the educationally disadvantaged high school students, the difference in average EOC test scores in DSSF versus other districts is about 0.4 points. Without the DSSF pilot, that difference would be more than 3 times as large, about 1.27 points. In contrast, it does not appear that DSSF was effective in reducing the gap between academically disadvantaged high school students and their more advantaged peers in the pilot districts. We base this conclusion on the finding that the average increase in test scores for all high school students in the DSSF districts was larger than the increase for academically disadvantaged students. Yet it also appears to be true that the gains of the academically disadvantaged students did not come at the expense of proficient students. Confidence in the validity of these estimates should be increased by the fact that evidence indicates that (a) a "fly-paper" effect did seem to occur, which resulted in additional spending in the pilot districts, (b) additional funding went into categories of expenditures that are plausibly related to improvements in student achievement (regular instruction, professional development, and instructional support), and (c) other reforms and the accountability pressures that were operating in the state were not disproportionately concentrated in the pilot districts, and therefore the threat to the validity of these estimates of the effects of the supplemental funding that could have arisen from other sources is diminished.

The positive effect of supplemental funding on high school student achievement was found for 16 of the most educationally disadvantaged districts in North Carolina, about 14% of the districts in the state. With the current interest in high school reform, and contrasted to highly prescriptive reforms that focus on changing processes within schools but do not provide increased funding for teachers in schools serving academically disadvantaged students, these results should be of interest. The intervention tested here requires only a modest investment of new funds, a minimal technical assistance capacity, and an index of educational advantage or disadvantage. Most other states could meet the requirements to implement a program like the DSSF intervention. By using an index of educational advantage like the one used to assign districts to the DSSF intervention, other states could both target funds to educationally needy districts and make their interventions rigorously evaluative via a RD design. However, the effects that were found in this study cannot be formally extrapolated beyond the most educationally disadvantaged districts. The estimates for the program effects presented here are local average effects or, stated another way, effects on the treated estimated at the cutoff. The estimates did not change to any meaningful degree when the estimates were limited to the 32 districts, the 16 above and 16 below, closest to the cutoff. Strictly speaking, the effects of supplemental funding can be extrapolated only to the most disadvantaged districts in a state, but within these districts, the effects are meaningful and important.

However, the results may provide clues to policymakers whose goal is to improve the educational achievement of academically disadvantaged students in a state. That is, the strategy of supplementing the state funds for the districts with the highest concentrations of the most educationally disadvantaged students produced benefits in terms of student achievement. Therefore, in states such as North Carolina it seems reasonable to expect that gains could be made in districts with high concentrations of educationally disadvantaged students by increasing state funds to those districts.

The supplemental funding program has specific elements that may have made it more effective. First, the DSSF funding was directed to districts. Providing funding to districts rather than school creates more flexibility in how the funds can be used. When supplemental funding goes directly to schools, it often is spent on pull-out programs, class size reduction, or other special programs that can be funded for the individual schools. At the district level, where the ability to control teachers' compensation resides,
there is flexibility to spend funds for bonuses or to increase teachers' salaries as well as for mentoring, professional development, and instructional supports that may directly affect the quality of teachers and teaching. Increasing teachers' compensation in districts with high concentrations of educationally disadvantaged students may allow these districts to more effectively compete to hire and retain effective teachers and improve the morale and effort of these teachers. The effects documented in this study were achieved through spending on a menu of "proven strategies" established by the state that focused on teacher recruitment, development, and retention, class size reduction, and individualized instruction for students. The districts spent the funds on regular instruction, mainly salaries, professional development, and instructional support. The strategy of providing supplemental funding to districts with guidelines that allowed for some local flexibility seems to have allowed educators and administrators to focus on specific problems.

In North Carolina or other states if they pursue this strategy, it seems most reasonable to focus funding on the most disadvantaged districts until evidence indicates that there are diminishing returns for the funding increases. It also seems reasonable to increase the number of districts eligible for funding by moving the cutoff for funds incrementally to include the group of districts that are just above the original cutoff in the distribution of additional supplemental funding. The evidence suggests not only that the most disadvantaged districts are likely to produce modest but positive effects from the increases but also that high concentrations of academically disadvantaged students in these districts may result in higher achievement for students who struggle the most. Unfortunately, the expansion of the DSSF strategy in North Carolina spread much smaller amounts of funding across all of the districts with few meaningful guidelines for how it could be used. However, the pilot districts have continued to receive nearly half of the funding, which may allow them to continue to supplement salaries and provide strategic instructional support and professional development.

Note

1. Two additional exams were given for the first time in 2004–2005, our first study year (American history, and civics and economics) but were excluded because the high schools had not had student scores in prior years and therefore had no time to expend funds to improve these scores or adjust their behaviors. A large qualitative study provided evidence that the adjustment process was under development in 2005–2006 in many of the state's high schools (Thompson, Brown, Cunningham, & Montrosse, 2008).

References


Targeted Funding for Educationally Disadvantaged Students


Authors

GARY T. HENRY is Duncan & Rebecca Kyle MacRae Professor of Public Policy at the University of North Carolina at Chapel Hill, 122 Abernethy Hall, Chapel Hill, NC 27599-3435; ghenry@unc.edu. He also directs the Carolina Institute for Public Policy at UNC Chapel Hill. His areas of specialty are education policy, evaluation, and quantitative research design and methods.

C. KEVIN FORTNER is a researcher at the Carolina Institute for Public Policy at the University of North Carolina at Chapel Hill, 117 Abernethy Hall, Chapel Hill, NC 27599-3435; cfortner@unc.edu. His research interests have focused on evaluation, policy analysis, and quantitative methods applied to K–12 schools.

CHARLES L. THOMPSON is LW King Professor in Education at East Carolina University, 205 Ragsdale Hall, Greenville, NC 27858-4353; THOMPSONCHAR@ecu.edu. His current research concerns the impact of teacher preparation routes and programs as well as state policy interventions on student achievement in North Carolina.