THE IMPACT OF COURT-ORDERED SCHOOL DISTRICT FINANCE REFORMS ON HIGH SCHOOL PROFICIENCY TEST SCORES IN NEW JERSEY’S ABBOTT DISTRICTS

A Thesis
Submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy
in Public Policy

By

Marnie B. Kaplan, M.S.T

Washington DC
April 15, 2011
THE IMPACT OF COURT-ORDERED SCHOOL DISTRICT FINANCE REFORMS ON HIGH SCHOOL PROFICIENCY TEST SCORES IN NEW JERSEY’S ABBOTT DISTRICTS

Marnie B. Kaplan, M.S.T.

Thesis Advisor: Laura LoGerfo, Ph.D

ABSTRACT

In 1997, the New Jersey State Supreme Court declared the state’s system of education funding unconstitutional and ordered the state to provide the poorest school districts—which were deemed the Abbott districts—with per-pupil funding equal to that of wealthy suburban districts. This project seeks to determine how the court mandated reforms are related to educational outcomes of New Jersey’s minority students. Specifically, this paper examines whether the increased funding led to more students scoring proficient on the High School Proficiency Assessment for students in Abbott districts as compared to students in districts which did not receive increased funding. The results suggest that the Abbott reforms specifically—and increased funding generally—are not linked to improved district-level performance on the High School Proficiency Assessment, which students must pass in order to graduate high school in New Jersey.
TABLE OF CONTENTS

Introduction.............................................................................................................. 1

Chapter I.................................................................................................................. 3

Background............................................................................................................ 3

History of School Finance Reform in New Jersey............................................... 3

Positive Relationships Between Expenditures and Performance.................. 7

No Relationship Between Expenditures and Performance............................ 10

Unclear Relationships Between Expenditures and Performance................... 10

Expenditures and High School Graduation and Dropout Rates...................... 12

Previous Evaluations of Abbott Reforms......................................................... 13

Research Hypothesis.............................................................................................. 16

Chapter II ............................................................................................................... 17

Data....................................................................................................................... 17

Sample................................................................................................................... 18

Method.................................................................................................................. 21

Results................................................................................................................... 23

Chapter III ............................................................................................................ 27

Discussion.............................................................................................................. 27

Tables..................................................................................................................... 29-32

References............................................................................................................. 33
Introduction

Disparities in academic achievement between low-income and high-income students and between minority and white students represent a pervasive education problem in the United States. In the aftermath of the landmark Supreme Court case to end racial segregation in American schools, *Brown vs. Board of Education*, integration was viewed as a mechanism for diminishing this disparity by allowing black students access to the same schools and resources as their white peers. In the 1970’s, as desegregation stalled, school finance reform was viewed as the primary mechanism for creating equal educational opportunities for racially and economically diverse students. Since then, school finance reform has been one of the dominant forces in state level educational reform; yet its effectiveness as a reform is questioned. Researchers debate whether more financial investment and increased expenditures reduce achievement gaps by race and socioeconomic status. Two separate meta-analyses of the link between resources and student success came to opposite conclusions regarding the relationship (Hanushek, 1986, 1996; Hedges, Laine and Greenwald, 1994). Hanushek (1996) asserts that increased funding does not reduce inequalities in student outcomes, whereas Hedges et al (1994) conclude that increased funding has a positive effect on student outcomes. Despite these unresolved debates, states have continued to alter their funding equations in an effort to close state-level achievement gaps. Important questions remain unanswered: Is this money well spent? Does increasing per-pupil funding provide low-income students with better educational opportunities?

The public school system of New Jersey provides an opportunity to examine the relationship between successive years of increased funding and students’ educational
outcomes. New Jersey was one of the first states to be embroiled in school finance litigation and one of the first states to attempt to reform inequalities in spending across school districts. Since 1997, thirty-one low-income urban districts in New Jersey have been provided with school funding to match that of their suburban neighbors, a move called ‘parity funding.’ An analysis of New Jersey thus provides the opportunity to examine the cumulative impact of multiple years of increased funding. In addition, studying New Jersey provides a lens on a geographically diverse state, one with a large number of cities, wealthy suburbs and rural communities.

Additionally, New Jersey has been the focus of a great deal of recent media attention as the Schott Foundation declared that New Jersey’s Abbott reforms had led to a closing of the achievement gap between white and black males (Yes We Can: The Schott 50 State Report on Black Males and Education). Additionally, Marc Zuckerberg, the founder of the popular social networking site Facebook, donated a million dollars to the Newark school system (an Abbott school district). Newark’s Mayor Cory Booker is attempting to match this donation with an additional million from private donors. Before more money is pumped into the Newark schools and other similar districts it is necessary to examine the effectiveness of the various Abbott reforms.
Chapter I. Background

History of School Finance Reform in New Jersey

New Jersey’s school finance reform began in 1970, when Jersey City sued the state on behalf of Kenneth Robinson, an 11-year-old growing up in a housing project. That suit, Robinson v. Cahill, asserted that the state’s heavy reliance on local property taxes to pay for schooling ensured that poor school districts could never spend as much as rich ones, even while taxing far more heavily (Robinson v. Cahill, 1973). The state Supreme Court ruled that such inequities violated the state constitution’s guarantee of a “thorough and efficient” educational system (Robinson v. Cahill, 1973). In response, the 1975 New Jersey Public Education Act moderated to some extent the per-pupil expenditure inequalities between school districts by increasing state aid to poorer school districts. In 1981, the Education Law Center, a non-profit organization, filed the state’s second school finance lawsuit Abbott v. Burke, which challenged the persistent inequalities in per pupil expenditures. At the time this case was brought, poor districts were spending about $2,400 per pupil while taxing at $1.65 per $100 assessed valuation, and rich districts were spending $3,000 per pupil while taxing at 97 cents per $100. The Education Law Center brought the case on behalf of 20 children who lived in Camden, East Orange, Irvington and Jersey City. The Education Law Center argued that the Public School Education Act of 1975 did not sufficiently ameliorate the educational disparities between poor and wealthy districts (Ollenschleger, 1091). After first remanding the case to an administrative law judge to develop a factual record in Abbott I, the court in Abbott II found for the plaintiffs, unanimously holding that the 1975 Act violated the New Jersey Constitution as applied to twenty-eight lower-wealth urban school districts (Abbott v.
Burke II, 1990). The court therefore ordered the legislature to amend the Act to ensure “substantially equivalent” per-pupil expenditures in property-rich and property-poor school districts (Abbott v. Burke II, 1990). The court also declared that supplemental programs be initiated in the urban districts to mitigate educational disadvantages (Abbott v. Burke II, 1990).

The court in this decision limited the Abbott reforms to a class of school districts identified as “poorer urban districts” or “special needs districts.” In 1997, these districts became known as “Abbott districts” (Abbott II, 1990). The Court identified four specific factors used to designate districts as “Abbott districts.” The district must: 1) be assigned to the lowest category on the New Jersey Department of Education’s District Factor Groups scale, i.e. one with the lowest socio-economic status; 2) display “evidence of substantive failure of thorough and efficient education” including “failure to achieve what the DOE considers passing levels of performance on the High School Proficient Assessment (HSPA)” ; 3) have a large percentage of disadvantaged students who need “an education beyond the norm”; and 4) have “excessive tax [for] municipal services” in the locality where the district is located (Abbott II, 1990). Using these factors, the Court in Abbott II identified 28 districts as Abbott districts. The Court also gave the New Jersey Legislature or the Commissioner of Education the authority to classify additional districts as Abbott districts based on these factors, which would then entitle these districts’ children to the Abbott programs and reforms (Abbott II, 1990). In 1998, the Legislature classified three additional districts, bringing the 2009 total of Abbott districts to 31.

Immediately after Abbott II was handed down, Governor Florio introduced a new funding bill into the state legislature and two weeks later both houses passed the Quality
Education Act of 1990 (QEA). The QEA provided for increased state aid to the Abbott districts, a phase out of state aid to wealthy districts, and a reduction in state funding for teachers’ pension plans. (Reed, 138). The QEA never went into effect as politicians responded to the suburban voters and the NJEA, New Jersey’s largest teacher organization two groups which aggressively opposed QEA. Politicians replaced QEA with QEA II which decreased the tax burden for state residents, provided for less of an increase in state aid to Abbott districts, and delayed shifting the costs of teachers’ pensions to local school districts (Ollenschleger 1097-1098).

In Abbott III, the New Jersey Supreme Court declared the Quality Education Act II unconstitutional both because it failed to ensure parity funding by the 1995-96 school year and because it did not provide supplemental programs to help disadvantaged Abbott students (Abbott III, 1994).

The most dramatic change in New Jersey school finance occurred in Abbott IV, in which the State Supreme Court ordered New Jersey to provide the poorest urban districts (the Abbott districts) with the same per pupil funding as the high wealth suburban districts (Abbott IV, 1997). The legislature complied by adding nearly $250 million of “parity” funding to raise the spending levels in the Abbott districts to the same levels as the state’s wealthiest suburban districts. A year later, in Abbott V, the court ordered specific education reforms to be implemented in Abbott districts including preschool for all three and four year olds, full-day kindergarten, whole school reform in elementary grades, full funding of new school construction, and supplemental educational services such as tutoring (Abbott V, 1998).

By 2007-2008 the Abbott Districts spent $1,300 per pupil more than the high
wealth districts and considerably more than both the low-wealth non-Abbott districts ($4,000) and the middle-wealth non-Abbott districts ($3,000) (Goertz and Weiss, 5) For ten years the Abbott districts received per pupil funding at the same level or higher than the state’s wealthiest suburban districts. In 2008, Chris Christie, New Jersey’s new governor, dramatically altered the state’s school financing when he implemented the School Finance Reform Act (SFRA). The SFRA replaced Abbott parity funding with a formula applicable to all districts that apportioned money based on school enrollment and the characteristics of students including family income, language ability and special academic needs. Under this formula, low- and middle-wealth non-Abbott districts benefited the most with an increase of $2,000 per pupil in regular and categorical education funds (Goertz and Weiss, 5-6). The Act included an “adjustment” aid provision that maintains current levels of funding (with a 2% increase) for the near future. Without this adjustment Abbott districts would have lost $1100 per pupil in revenues (Goertz and Weiss 6). In May 2009, the New Jersey Supreme Court ruled in Abbott XX that SFRA was constitutional and may be applied to the Abbott districts subject to the state fully funding the SFRA formula annually, conducting a review of the law after three years of implementation, and adjusting the formula’s adequacy amount, weights, and other elements based on a required review.

School finance litigation is not restricted to New Jersey. Over the past thirty years, 45 of the 50 states have been involved in school finance litigation (Hunter, 2005). In 28 of these cases, the court has ruled for the plaintiffs, children living in low-income communities. With this groundswell of support for such parity funding initiatives, these states should query whether the reform actually works. Therefore, analyzing the
relationship between school finance reform and academic outcomes such as graduation rates in New Jersey will advance the discussion of the general value of school finance reform.

**Positive Relationships Between Expenditures and Performance**

Motivating school finance litigation is the assumption that inequalities in educational opportunities can be attributed to inequalities in spending. Some researchers have found (Hedges et al., 1994; Krueger, 1998, 2002) that there is a systematic relationship between school expenditures and student performance. Hedges, Laine and Greenwald (1994) looked at the aggregated data of 60 existing research studies related to the input-output relationship for schools. Their analysis suggests that a broad range of resources are positively related to student outcomes, with effect sizes large enough to suggest that moderate increases in spending may be associated with significant increases in achievement. Specifically they find that a $500—or roughly ten percent—increase in average spending per pupil increases district-level student achievement by 0.07 standard deviations, a small but important margin. Among the specific educational resources money can buy, a number of practices have relatively strong positive relationships to the achievement of disadvantaged students; these include universal preschool, class-size reduction and whole school reform. Barnett (2002) demonstrated that intensive, high-quality preschool programs can close much of the early achievement gap for lower income children. For example, an evaluation of the Abbott Preschool Program (Frede et al., 2009) showed positive links between preschool attendance and children’s learning in language, literacy, and math through the end of second grade. The randomized
Tennessee class-size experiment provided strong evidence for the argument that funding matters, by showing class size reduction increased standardized test scores by about 4 percentile points in one year with stronger benefits for minority and low-income students (Krueger, 1997). In addition, a randomized experiment of a whole school reform model, Success for All, showed positive gains in literacy compared to the control group after two years in the program, with those subject to treatment demonstrating a six-month learning advantage on phonological decoding (Borman, Slavin & Cheung, 2005). While these studies do not examine finance reform specifically, financial resources allow schools to hire more teachers which facilitates smaller class size, invest in whole reform models and fund preschool programs. In fact, these particular education practices were implemented as part of the Abbott remedies.

Some analyses of school finance reform also support the conclusion that increased expenditures positively affect student achievement. In a cross-sectional study of 12 states with school finance reform, Card and Payne (2002) found that equalization of spending is tied to a narrowing of students’ SAT scores across family backgrounds. Their estimates imply that the spending equalizations that followed twelve states’ court rulings that the funding system was unconstitutional during the 1980s closed the gap in average SAT scores between children with highly-educated and poorly-educated parents by about eight points, or roughly five percent. Card and Payne (2002) also found evidence that the equalization of spending across districts narrows the gap in SAT participation between higher and lower socioeconomic groups. Since taking the SAT shows an intention to attend college, this finding suggests that the higher incidence of SAT-takers in low-income districts supports the assertion that increased funding leads to increased student
outcomes and overall better quality of schooling. However, Card and Payne’s use of the SAT is problematic as it introduces selection bias. While the SAT can be used to compare achievement levels across states, it is worth noting that students who take the SAT differ from the general population of high school students. Using two national data sets (NLS-72 and the 1992 wave of NELS:88) Downes and Figlio (1997) examine the mean test performance of high school seniors in districts with and without school finance reform, finding that court-mandated school finance reforms (which focus on utilizing equalization formulas) have led on average to increased student performance. But these results are sensitive to the inclusion or exclusion of California, which will be discussed in subsequent section at length.

Despite these studies that suggest benefits to academic outcomes from parity funding reforms, a study of Vermont by Downes (2004) offers counter evidence. In Vermont, gaps in performance between high spending and low spending districts and between high wealth and low wealth districts have declined after the implementation of Vermont’s school finance reform, Act 60 (Downes, 2004). Only small improvements emerge in the test performance of fourth and eighth graders in disadvantaged school districts that spent less money per pupil and had less per property wealth prior to the reform.

In sum, one side of the school finance debate enjoys a solid set of empirical evidence that higher expenditures has a positive impact on student achievement, but all of this evidence comes from meta analyses of a variety of states which could confound the results of individual state school finance reform. This possibility is illustrated by the results of Downes analysis of Vermont post Act 60.
No Relationships Between Expenditures and Performance

Despite the research just reviewed which shows benefits of higher expenditures, research by Hanushek (1986, 1989, 2003; Rivkin et al., 2005; Hanushek, et al., 1999) failed to find a relationship between increased expenditures on public education and educational performance. Surveying the same existing studies used by Hedges, Laine and Greenwald (1994), Hanushek concluded that there is no evidence that teacher-pupil ratios, teacher education, teacher pay, or spending per-pupil are consistently related to improving school achievement. Hanushek’s studies imply that equalization of funding should not be expected to be an effective policy for promoting greater equality of outcomes. He summarizes his case most dramatically by pointing out that student performance on achievement tests in the United States has been flat for 30 years while real expenditures per pupil have increased threefold. As Coate and Vanderhoff (1999) write: “Given the amount of litigation over inequalities in public school expenditures, it is surprising that the literature on public school expenditures and student performance is not supportive of a strong relationship between the two.”

Unclear Relationships Between Expenditures and Performance

Many of the analyses of state-level school finance reform have revealed an unclear relationship between expenditures and performance. In his analysis of California before and after Serrano v. Priest – California’s school finance court case-- Downes (1992) found that inequalities in per-pupil expenditures decreased but inequalities in student achievement did not. While this study is viewed by some as a state-level
illustration of Hanushek’s conclusion that increased funding does not yield increased educational outcomes, some caution must be used in drawing conclusions from school finance in California. First, California’s severe limits imposed on local control over spending have not been duplicated in any other states. In fact, the *Serrano v. Priest* decision is credited with keeping the state’s per-pupil spending lower than would have been the case if court-ordered finance reform had not taken place. (McCombs & Carroll, 2005). Second, the population of students served by California schools changed more dramatically than the population of any other state in the nation. From 1986 to 1997 (Downes reviewed 1976-1977 and 1985-1986), the percent of the California public school student population identified as minority increased from 46.3% to 61.2%. Nationally, the percent of minority students in K-12 grew far more slowly, from 29.6% to 36.5% (Downes 2003). Thus, California’s results are not generalizable to the nation as a whole and provide poor evidence about the relationship between resources and achievement.

Kentucky is another state that was galvanized to reform its funding structure. In 1990, the Kentucky Education Act (KERA) was passed in response to a ruling the previous year by the Kentucky Supreme Court that the commonwealth’s education system was inequitable and therefore unconstitutional. A primary aim of this act was to reduce economic disparities between schools. After its passage, districts began to receive a predetermined amount of money based on a per-student system. To examine the relationship between the passage of the Kentucky Education Act (KERA) and student achievement, Clark (2003) used student level data pre- and post-reform to conduct a difference-in-difference analysis using Tennessee and other neighboring states. The
results indicated that the passage of Kentucky Education Act (KERA) resulted in a small increase in black students’ ACT scores, but no statistically significant increase in black students’ NAEP test scores. She concludes that KERA has been largely unsuccessful in raising student achievement and in narrowing the gap in student achievement across rich and poor districts. Flanagan and Murray (2002) also examined student achievement in Kentucky after the implementation of KERA. From a regression of the relationship between the change of test scores and the change in per pupil spending, Flanagan and Murray’s model suggests that a one dollar increase in per pupil spending increases test scores by .01 percentage points. Thus, the conflicting results make it difficult to conclude whether KERA shows a relationship between expenditures and student achievement.

In 1994, Michigan altered its school finance system with the legislature-led Proposal A, which significantly increased state aid to the lowest spending school districts. Looking at finance reform in Michigan, Roy (2004) found significant gains in achievement in the poorest districts as measured by success on state tests. However, he found that these improvements did not spill over to other areas, like participation and performance in ACT or on the NAEP.

With such variability in impact across state finance reforms, the state-level analyses of California, Kentucky and Michigan imply that more research is necessary to understand the relationship between expenditures and student outcomes.

**Expenditures and High School Graduation and Dropout Rates**

High school graduation is a very important predictor of a young person’s life prospects. In 2004, high school dropouts earned only 37 cents for each dollar earned by individuals with more education (Rouse 2005). Rouse thus concludes: “High school
graduation has been a necessary (but not sufficient) pre-requisite for making it in America.” High school graduation rates are therefore an important measure of our public school system.

The relationship between per-pupil expenditures and graduation rates has received much less focus than the relationship between per-pupil expenditures and student achievement. One of the only evaluations was conducted by Hoxby (1998) to evaluate the relationship between state level school finance reform and the high school drop out rate. In her cross-sectional analysis of California, New Jersey, New Mexico, West Virginia, Delaware, Utah, Wyoming, Connecticut, Kansas and Kentucky, Hoxby finds that leveling up equalization schemes, i.e. those that increase funding in lower income districts, decrease the high school dropout rate slightly. Duncombe and Johnston’s (2002) analysis of Kansas before and after school finance reform suggests some recent improvement in dropout rates in high poverty districts in Kansas pre- and post-finance reform, although they also find increased dispersion in dropout rates when comparing pre-and post-finance reform data. Thus, the existing research suggests that school finance reform may not play a significant role in reducing drop out rates. As for high school graduation rates, while research studies evaluate the relationship between a state’s per-pupil funding and the statewide graduation rate (in other words, which states get the best return on their education investment), none of this work adjusts for differences in school district characteristics to isolate the relationship of school finance reform to graduation rate.

**Previous Evaluations of Abbott Reforms**

From the scant research on graduation rates and expenditures it is not clear what
the relationship may be between Abbott funding and assessments linked to graduation at the district level in New Jersey. Previous evaluations of the Abbott reform have been limited in scope, typically focus on younger student outcomes, and have not looked directly at the relationship between expenditures and likelihood of successful high school outcomes. Ritter and Lauver (2003) focused mainly on whether the reforms led to equalization in funding and only briefly examined the student performance of fourth graders in districts of different funding levels on a state-wide elementary school proficiency test. Their study reveals that students in the Abbott districts continue to trail behind their high-wealth, middle-wealth, and poor non-Abbott counterparts on the fourth grade exam; but Ritter and Lauver fail to examine the change over time to see if this gap is smaller than in previous years. Goertz and Weiss (2009) conclude that the Abbott reforms led to a significant closure in the achievement gap between the Abbott districts and the rest of the NJ districts in fourth grade and a small closure in the achievement gap between the Abbott districts and the rest of the NJ districts in eighth grade. However, these conclusions are based on only descriptive statistics. Coate and Vanderhoff (1999) look specifically at high school achievement as measured by high school proficiency test scores between 1988-1989 and 1994-1995 and find no relationship of expenditures on student achievement. While this is very similar to the research question driving this study, they examine results prior to the ruling for funding parity, which began in 1997. Therefore, the question of whether parity funding leads to increased scores on a test that students must pass to graduate remains unanswered. Additionally, their analysis does not measure the cumulative impact of a large number of years of increased funding.

The sheer amount of state money poured into the Abbott districts – a total of
$37.7 billion since 1998—has made them among the highest spending public school districts in the country (Hu, New York Times 2/12/2008). Critics of Abbott funding assert that districts in rural and suburban communities serve families who are just as poor as those in the Abbott districts and need the funds to serve these disadvantaged populations (Ritter and Lauver, 582). They assert that it should be no surprise that a remedy aimed at 31 of the state’s over 600 districts did not solve the systemic problem with funding inequities. In fact, in 2000, 17 rural districts sued to become Abbott districts alleging they meet all the Abbott factors and are unable to provide their students a constitutional education without the Abbott remedies. Chris Christie, New Jersey’s new governor, seems to have listened to critics who asserted that funding inequities would persist until the state lawmakers are willing to develop a remedy that is comprehensive. In 2008, Governor Christie dramatically altered New Jersey’s school financing scheme – apportioning money to school districts based on enrollment and the specific characteristics of their students, i.e. family income, language ability and special academic needs. In 2008, according to the New Jersey State Department of Education, 56.6 percent of state school aid went to the 31 Abbott school districts; in fiscal year 2009, the first year of the new school-funding formula to end the Abbott designation and allow spending to follow the neediest children, 55.1 percent of aid went to the former Abbott districts. While the focus in New Jersey has shifted away from parity funding for Abbott districts, the Abbott districts are still receiving a great deal of funding.

While New Jersey citizens continue to argue over how the state’s money is best allocated, research on the relationship between expenditures and student achievement is mixed. Hedges, Laine and Greenwald and Card and Payne offer evidence that increased
expenditures can positively impact student achievement. Hanushek concludes that there is no evidence that spending per-pupil has a consistent effect on improving school achievement. The state level analyzes conducted in California, Kentucky and Michigan do little to settle these contradictory conclusions.

The relationship of per-pupil expenditures to an assessment linked to graduation is neglected in this research, and few studies use a regression approach to investigate this potential relationship. Overall, there is not enough conclusive research to support the fact that many states have embraced school finance reform and certainly not enough evidence to support the argument that increased funding will lead to a “thorough and efficient” education. This paper seeks to evaluate how school district per-pupil funding relates to percent of students passing a high school proficiency assessment at the district level in a state with court-mandated parity funding for low-income urban school districts.

**Research Hypothesis**

The research question at the focus of this work: How does school district per-pupil funding relate to high school proficiency rates in a state with court-mandated parity funding for low-income urban school districts? Specifically, does New Jersey’s parity funding lead to an increased HSPA pass rate for students in districts that receive more per-pupil funding? I hypothesize that parity funding does not increase HSPA proficiency rates in the Abbott districts relative to the HSPA proficiency rates in other less-wealthy, non-Abbott districts (most of which are rural).
Chapter II. Study

Data

The primary data source for this study is the New Jersey Department of Education (NJDOE), specifically New Jersey’s High School Proficiency Exam (HSPA) data, enrollment data and mean SAT scores from the New Jersey School Report Card for the 2007-2008 school year and per-pupil spending data from the 2007 New Jersey Comparative Spending Guide. Because per pupil spending data is only available at the district level, district-level HSPA rates and enrollment data are analyzed. The analytic sample is restricted to public school districts that are not public charter, vocational or special service districts, which enroll children with special education needs.

New Jersey tests 11th graders each year on the HSPA which covers Language Arts Literacy, and Mathematics. This test was administered for the first time in Spring 2002. In the Spring of 2008 the HSPA was administered to 99,570 first-time 11th graders. Satisfactory HSPA Performance in Mathematics and Language Arts Literacy is a requirement for a high school diploma. The HSPA scores are reported as scale scores in each of the content areas. The scores range from 100-109 (Partially Proficient), 200-249 (Proficient) and 250-300 (Advanced Proficient). The Literacy assessment includes multiple-choice and open-ended answers including an oral component and an editing assignment. For the editing assignment students are required to revise a piece of student text that has errors in sentence structure, usage, mechanics and content. For the oral component students are given 60 minutes to prepare a speech on a given topic. Student’s responses to open-ended items are scored using a 0-to 4-point or a 0-to 6-point rubric. The Math assessment includes 26-31 multiple-choice questions, 4 short constructed-
response questions and 7 open-ended questions. For the short constructed-response items students are required to write down algebraic expressions, equations and simple graphs in response to a question. For the open-ended items students are required to construct their own written or graphical responses.

**Sample**

This project analyzes a one-year snapshot of data, comparing HSPA proficiency rates across the 266 school districts in New Jersey that contain high schools. The year of analysis is 2007-2008, the last year of parity funding. While it would have been preferable to look at graduation rates in 2007-2008 or at the changed graduation rate over the entire time period of Abbott reforms, the graduation rates documented in the 1997 and 2008 New Jersey Report Cards include imprecise data due to the fact that there was little accountability for graduation rate data during this time period. The intention of this snapshot is to determine how increasing per-pupil funding is related to high school outcomes for students who benefited from ten years of increased funding.

In order to isolate the relationship between increased per-pupil funding and the Abbott reforms, I account for variation in districts’ academic and demographic characteristics. Specifically, I adjust for differences in district achievement level using mean SAT math and verbal scores. Performance on the High School Proficiency Test may be related to district funding but is also likely related to how well district students perform on other standardized tests, such as the SAT. Inclusion of the mean Math and Verbal SAT scores at the district-level also addresses the fact that some districts may

---

1 There are 604 total school districts in New Jersey.
have instituted successful academic reforms previous to the year in question. Differences in student composition across districts are accounted for, at least in part, by including the percentage of minority students in each district. Including district percentages of black, Hispanic and Asian students accounts for findings that show black and Hispanic students are less likely to graduate high school and Asian students are more likely to graduate high school than white students. Each of these percentages is represented by a continuous variable. Differences in district size are also accounted for in the model. Including the number of enrolled high school students in each district accounts for research that shows smaller districts are often better able to meet their students’ academic needs. This is represented by a continuous variable.

To capture per-pupil funding, I divided the districts into quartiles based on their 2007-2008 per-pupil funding and reviewed the descriptive statistics. My goal was to evaluate how districts with different levels of wealth vary in terms of demographics and academic characteristics. Table 1 contains basic descriptive statistics on key demographic variables at the district level for districts whose per-pupil funding falls into the 25th, 50th, 75th or 99th percentile derived from multiple Bonferroni tests. Looking across these quartiles, the only significant difference in terms of percentage of black students is between the second quartile (i.e. the quartile representing districts with the second lowest per-pupil funding) and the fourth quartile (the districts with the highest per-pupil funding). The districts with the highest per-pupil funding (quartile 4) enroll significantly higher percentages of black students than districts with the second lowest per-pupil funding (quartile 2). Specifically, the percent of black students in quartile two is 12.2%, whereas the percent of black students in quartile four is 21.8%. This difference
most likely emerges because the Abbott districts all fall into Quartile 4, and the Abbott
districts, although selected by funding differentials, comprise cities with higher
concentrations of minority families, such as Newark and Camden. For example,
according to the recent Census the population of the state of New Jersey is 13.6% black
and 15.9% Hispanic while the population of Newark is 51.8% black, and 31.8% Hispanic
and the population of Camden is 49.9% black and 42.1% Hispanic. The fact that there is
no significant difference between Quartile 1 (the lowest funded districts) and Quartile 4
(the highest funded districts) in terms of percentage of black students is likely explained
by the fact that the percentage of black students in Quartile 4 is decreased due to the fact
that it is comprised of districts which are almost entirely black and districts which are
almost entirely white. In addition, the low funded districts in Quartile 1 still have some
minority concentrations. The districts in Quartile 1 are more likely to be similar in racial
composition whereas those in Quartile 4 are more likely to have extremes in racial
breakdown.

Similar to the findings for district percentages of black students, Quartile 4 includes
significantly higher percentages of Hispanic students than Quartile 3 and Quartile 2
Quartile 2 includes a mean of 12.18% Hispanic students, Quartile 3 has a mean of
12.14% Hispanic students whereas Quartile 4 has a mean of 21.81% Hispanic students.
This once again results from the fact that the high concentration of Hispanic students in
the Abbott districts heavily skews the average in Quartile 4. The four quartiles enroll
roughly similar percentages of Asian students.

When comparing the size of the school districts within each quartile, statistically
significant differences are found among quartiles 2 and 4, and quartiles 3 and 4. The
difference in size is largely a result of the much larger number of students enrolled in schools in Quartile 4, which once again reflects the fact that the Abbott districts (which contain large inner city high schools) all fall into this category.

**Method**

Ordinary least squares regression (OLS) is used to evaluate whether an increase in per-pupil funding is linked to an increase or decrease in high school proficiency test scores. OLS is a method for estimating the unknown parameters in a linear regression model. This method minimizes the sum of squared vertical distances between the observed responses in the dataset, and the responses predicted by the linear approximation. In other words, the β co-efficients, which are estimated through OLS estimates, generate the best-fit line through the data. OLS estimators have nice properties under three key assumptions: on average the predictions are correct (E=0); the variables are independently and identically distributed; and there are no huge outliers in the data.

An OLS regression equation expresses the unique linear contribution of each X-variable to the Y-variable. In the hierarchical modeling I have set up, each β co-efficient will explain whether the variable has a positive or negative relationship with percent of proficient students and the magnitude of that correlation. Each of the three regressions is carried out in order to test the null hypothesis (H₀) that there is no relationship between the variables of interest and the dependent variable—percent of student proficient on the HSPA. For example, in Model 1, OLS regression is used to test the null hypothesis (H₀) that there is no relationship between each demographic and academic characteristics of a district and the percent of students proficient on the HSPA exam. The alternate
hypothesis ($H_1$) is: there is a relationship between each demographic and academic characteristics of a district and the percent of students proficient on the HSPA exam.

Another valuable result of an OLS regression is being able to determine which variables help the model, in other words increase the R-squared. R-squared is a measure of association. It represents the percent of the variance in the values of the Y that can be explained by the X values. R-squared varies from a low of 0.0 (none of the variance is explained), to a high of 1.0 (all of the variance is explained). The models used in this study will determine if the variables per-pupil funding and Abbott explained more of the variance in the percent of students proficient on the HSPA exam.

Since the population parameters (population mean and population standard deviation) for the original population in the sample are known the z-test is used instead of a t-test.

In Model 1, social and academic background is regressed on district percent passing the HSPA. Model 1 is:

$$ P_{\text{Proficient}}_{dt} = \beta_0 + \beta_1 \text{BLACK}_{dt} + \beta_2 \text{HISP} \_dt + \beta_3 \text{ASIAN} \_dt + \beta_4 \text{SIZE} \_dt + \mu $$

Per-pupil funding--the variable of interest--is not included until the second iteration of the model. Adding the variable of interest in a separate step isolates the explanatory contribution of per-pupil funding by isolating the change to the R-squared uniquely related to per-pupil funding. In other words, this hierarchical model or step approach tests if per-pupil funding accounts for any variation in percent passing the HSPA beyond that explained by differences in district demographic and academic characteristics.

In Model 2, per-pupil funding and social and academic background is regressed on percent passing the HSPA. Model 2 is represented by the equation:
This model estimates the percent increase in HSPA scores that results from a one-dollar increase in funding. In order to see if Abbott reforms have any relationship with student achievement outcomes separate from changes in spending levels, a dummy variable for Abbott is added in Model 3. The equation for model three is:

$$P_{Proficient}^{dt} = \beta_0 + \beta_1 BLACK_{dt} + \beta_2 HISP_{dt} + \beta_3 ASIAN_{dt} + \beta_4 SIZE_{dt} + \beta_5 PP + \beta_6 ABBOTT + \mu$$

This model examines if the Abbott interventions have any relationship with academic outcomes independent of increased funding. The goal of the third model is to see if adding Abbott will help the model, in other words if adding Abbott to the model increases the R-squared. Each of these models is run twice, once to examine the HSPA in language (Model 1A, 2A, 3A) and once to examine the HSPA in math (Model 1B, 2B, 3B).

**Results**

The findings presented in Table 3 indicate that districts’ per-pupil expenditures are negatively related to district average achievement on the High School Proficiency Assessment. This relationship is small, yet significant with a one-dollar increase in per-pupil funding associated with a 0.001 (p<.05) decrease in the percentage of students proficient on the HSPA language and a 0.001 (p<.05) decrease in percentage of students proficient on the HSPA math exam. Admittedly these are very small decreases due to per-pupil expenditures, suggesting that monetary inputs are not providing the intended
outputs. Although the relationship is significant, the relationship is weak, suggesting no substantive or meaningful difference in test scores due to per-pupil funding.

The R-squared is slightly higher in Models 2A and 2B (.49 and .44) as compared to the R-squared in Models 1A and 1B (.42 and .41). Therefore, adding per-pupil funding increases the predictive power of the equation. But, compared to expenditures, the demographic variables have far stronger relationships with the outcome. Per-pupil funding has a lower magnitude of relationship to either outcome than every other variable except for total number of enrolled students. Adding per-pupil funding to the model with HSPA language proficiency as the outcome lessened the magnitude of the negative coefficients for percentage of black students and percentage of Hispanic students and the magnitude of the positive coefficient of percent of Asian students. This suggests that the underlying relationship between minority percentage and district performance may be due to the differences in per-pupil funding. Therefore, per-pupil funding may be acting as a mediating variable, not a true mediator since the demographic variables are still significant.

Results in Table 4 show that the relationship of per-pupil funding to HSPA performance is marginally smaller in a model that includes whether a district is within the Abbott group. In this model, a one-dollar increase in per-pupil funding leads to a .00055 decrease in the percentage of students proficient on the HSPA Language examination as compared to a per-pupil funding leading to a .001 decrease in Model 1A (which does not account for per-pupil funding); a one-dollar increase in per-pupil funding leads to and a .00075 (p<.01) decrease in the percentage of students proficient on the HSPA Math examination as compared to a .001 decrease in the percentage of students proficient on
the HSPA Math examination in Model 1B (which does not account for per-pupil funding). Accounting for districts’ designations as Abbott districts decreases the relationship between per-pupil funding and HSPA scores by a small margin. This may imply that the Abbott interventions held a slight benefit to districts in addition to merely increasing financial resources.

The R-squared is the highest in models 3A and 3B (.53 and .55). Thus, adding Abbott to the model increases the predictive power of the equation. This illustrates that the Abbott interventions have a relationship with HSPA proficiency independent of the relationship between per-pupil funding and HSPA proficiency. In fact, Abbott is negatively related to district average achievement on the High School Proficiency Assessment. This relationship is large and significant with Abbott district status associated with a 7.93 (p<.05) decrease in the percentage of students proficient on the HSPA verbal exam and a 9.34 (p<.05) decrease in the percentage of students proficient on the HSPA math exam. This suggests that the Abbott reforms have not made a unique contribution to students’ achievement at the high school level. In addition, this suggests that Abbot reforms have not affected a high-needs student’s likelihood of graduating high school (since this assessment is linked to graduation).

Both of these tables confirm the original hypothesis, that parity funding does not increase HSPA proficiency rates in the Abbott districts relative to the HSPA proficiency rates in other less-wealthy, non-Abbott districts. These results suggest that increased per-pupil funding is not associated with an increase a district’s likelihood of increasing the percentage of students who meet New Jersey’s standard for high school proficiency. In other words, New Jersey did not increase the academic outcomes of students in high-need
districts by providing these districts with more funding per student.

In each model, districts with greater percentages of black students or of Hispanic students had lower mean scores on the HSPA exam. Only Model 2B indicates that districts with higher percentages of Asian students fared significantly worse than districts with lower percentages of Asian students on the HSPA math exam. This result conflicts with existing literature which shows that districts with higher percentage of Asian students have better overall reports of student achievement. It is unclear why the coefficient for percentage of Asian students turns negative in the model which examines Math proficiency and accounts for differences in per-pupil funding. One potential explanation is that this model poor Asian immigrants in low-income districts.
Chapter III. Discussion:

The results suggest that the Abbott reforms specifically—and increased funding generally—are not linked to improved district performance on the High School Proficiency Assessment, which students must pass in order to graduate high school in New Jersey. The results imply that increasing district-level funding is not sufficient to increase a student’s likelihood to graduate high school. This should advise state policymakers that increased spending is not a panacea and that it is imperative to consider what kind of educational reforms are tied to that spending. Additionally, in a time of fiscal responsibility and tight budgets, this stands as potential evidence that decreasing funding is not directly tied to decreasing student achievement or likelihood of graduation.

This paper hardly gives a definitive answer to the efficacy of the Abbott reforms on student outcomes in New Jersey or the relationship between per-pupil funding and students’ high school outcomes. Some caution should be taken with interpreting these results. First, this study was limited to district-level analysis, which means that important variation within districts and schools related to student outcomes may not be captured in these models. Smaller units of analysis, either school- or student-level would allow greater precision in estimating the importance of per-pupil spending and Abbott reforms on student educational outcomes. In addition, it would be beneficial for further research that can examine this question over time. This one-year snapshot does not clearly elucidate if increased funding increased the rate of low-income students who met New Jersey’s graduation requirements in the years following parity funding.

While there is a need for further research, this paper suggests that the $37.7 billion that has been pumped into the Abbott districts has not been effective in providing New
Jersey’s highest need students with the “thorough and efficient” education campaigned for during countless years of litigation. The students in the Abbott districts benefited from ten years of increased funding and a series of specific interventions but neither of these factors increased the total number of students that were able to pass a test required for graduation. While the relationship between expenditures and student achievement is still unclear, this study reinforces the proposition that how education funding is spent matters. As districts and state courts continue to grapple with how to improve student outcomes in low-income districts, the findings of this paper suggest that court-implemented finance reforms alone are not the answer.
Table 1: Comparing Academic and Demographic Characteristics among Quartiles based on District Per-Pupil Funding (n=261)

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>PP Funding Quartile 1</th>
<th>PP Funding Quartile 2</th>
<th>PP Funding Quartile 3</th>
<th>PP Funding Quartile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Asian</td>
<td>6.7</td>
<td>6.48</td>
<td>8.87</td>
<td>5.31</td>
</tr>
<tr>
<td>% Black</td>
<td>12.8</td>
<td>10.34*</td>
<td>12.09</td>
<td>20.08</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>15.4</td>
<td>12.18 **</td>
<td>12.14**</td>
<td>21.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V SAT mean</td>
<td>477</td>
<td>496</td>
<td>501</td>
<td>470</td>
</tr>
<tr>
<td>M SAT mean</td>
<td>496</td>
<td>510</td>
<td>523†</td>
<td>486</td>
</tr>
</tbody>
</table>

| Enrollment                  | 1482                  | 1170**                | 1141**                | 1800                  |

* statistically significant compared to Quartile 4 at the 1% level  
** statistically significant compared to Quartile 4 at the 5% level  
† statistically significant compared to Quartile 1 at the 5% level
Table 2: Regression of Demographics and District-Level Academic Data on HSPA Proficiency Rates (n=261)

<table>
<thead>
<tr>
<th></th>
<th>Percentage of Students Proficient on High School Proficiency Language Examination</th>
<th>Percentage of Students Proficient on High School Proficiency Math Examination</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Black</td>
<td>-29.33** (2.75)</td>
<td>-44.51** (3.51)</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>-17.26** (2.95)</td>
<td>-20.66** (3.48)</td>
</tr>
<tr>
<td>% Asian</td>
<td>3.68 (5.11)</td>
<td>-9.41 (6.64)</td>
</tr>
<tr>
<td>Mean Verbal SAT Score</td>
<td>-0.06** (.014)</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean Math SAT Score</td>
<td>N/A</td>
<td>-0.12** (0.02)</td>
</tr>
<tr>
<td># Enrolled</td>
<td>0.00* (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>R²</td>
<td>0.42</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Standard Error in parentheses; * significant at 5%, ** significant at 1%
Table 3: Relationship between Per-Pupil Funding and Percent of Students Scoring Proficient on the HSPA (n=261 districts)

<table>
<thead>
<tr>
<th></th>
<th>Percentage of Students Proficient on High School Proficiency Language Examination</th>
<th>Percentage of Students Proficient on High School Proficiency Math Examination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per-Pupil Funding</td>
<td>0.00** (0.00)</td>
<td>0.00** (0.00)</td>
</tr>
<tr>
<td>% Black</td>
<td>-26.03** (2.67)</td>
<td>-40.02** (3.39)</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>-14.29** (2.85)</td>
<td>-16.95** (3.33)</td>
</tr>
<tr>
<td>% Asian</td>
<td>1.72 (4.85)</td>
<td>-13.01* (6.27)</td>
</tr>
<tr>
<td>Mean SAT Score</td>
<td>-0.05** (0.01)</td>
<td>-0.11** (0.02)</td>
</tr>
<tr>
<td># Enrolled</td>
<td>0.00* (0.00)</td>
<td>0.00 (1.15)</td>
</tr>
<tr>
<td>R²</td>
<td>0.49</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Standard Error in parentheses; * significant at 5%, ** significant at 1%
Table 4: Relationship between Per-Pupil Funding and Percent of Students Scoring Proficient on the HSPA, Controlling for Abbott District Designation (n=261 districts)

<table>
<thead>
<tr>
<th></th>
<th>Percentage of Students Proficient on High School Proficiency Language Examination</th>
<th>Percentage of Students Proficient on High School Proficiency Math Examination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per-Pupil Funding</td>
<td>0.00** (0.00)</td>
<td>-0.00** (0.00)</td>
</tr>
<tr>
<td>Abbott</td>
<td>-7.93** (1.62)</td>
<td>-9.34** (1.91)</td>
</tr>
<tr>
<td>% Black</td>
<td>-24.64 (2.57)</td>
<td>-38.29** (3.27)</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>-11.80** (2.77)</td>
<td>-13.46** (3.27)</td>
</tr>
<tr>
<td>% Asian</td>
<td>-0.87 (4.68)</td>
<td>-15.14* (6.02)</td>
</tr>
<tr>
<td>Mean Verbal SAT Score</td>
<td>-0.06** (0.01)</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean Math SAT Score</td>
<td>N/A</td>
<td>-0.13** (0.02)</td>
</tr>
<tr>
<td># Enrolled</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.53</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Standard Error in parenthesis; * p<0.05 ** p<.01
References


Goertz, Margaret E. & Weiss, Michael (Nov. 2009). Assessing Success in School Finance Litigation: the Case of New Jersey, Education, Equity and the Law, No. 1


http://nieer.org/resources/research/multistate/nj.pdf


