2013

Have Federal Sanctions Helped Failing Schools? The Impact of No Child Left Behind in Texas

Ernest W. Hayhurst

CLAREMONT McKENNA COLLEGE

HAVE FEDERAL SANCTIONS HELPED FAILING SCHOOLS? THE IMPACT OF NO CHILD LEFT BEHIND IN TEXAS

SUBMITTED TO

PROFESSOR SERKAN OZBEKLIK

AND

DEAN GREGORY HESS

BY

ERNEST WAYNE HAYHURST

FOR

SENIOR THESIS

SPRING 2013

APRIL 29, 2013

Table of Contents

Acknowledgements

I would first like to thank Professor Ozbeklik for his guidance this semester; for without him, this paper would have been impossible. By writing this thesis I have learned so much about econometric modeling and various research methods, so for that Professor, I thank you. I would also like to thank all of my friends for joining me in the most incredible four years of my life. While the thought of leaving Claremont is a sad one, I cannot wait to see what all of you accomplish. Lastly, and of course not least, I'd like to thank my family for their unending support.

Abstract

This paper will assess the effectiveness of the No Child Left Behind Act (NCLB) in the state of Texas. In order to do this, we examine how students' performance levels from failing schools respond to sanctions imposed by the NCLB accountability system. Additionally, we explore achievement gap trends between white and minority students that attend these failing schools. By taking advantage of campus and year fixed effects, as well as controlling for student demographic characteristics, we find that sanctions employed by NCLB have had a statistically significant positive impact on academic achievement gains for all students. However, our results also indicate that these sanctions have effectively widened the achievement gaps between the white and minority students they affect. Given that the federal government spends upwards of 14 billion dollars per year to fund NCLB, this paper offers new insight to an economically important issue that is relevant to all citizens of the United States.

I. Introduction

In early 2001, President George W. Bush delivered an unprecedented plan for nationwide education reform to Congress. In this document, the President asserted that "the federal government must be wise enough to give states and school districts more authority and freedom. And it must be strong enough to require proven performance in return"[1]. Titled the No Child Left Behind Act (NCLB), it sought to amend the inefficient Elementary and Secondary Education Act of 1965[2] by installing a strict accountability system into America's educational framework. The reason for such measures was justified; at the time, the achievement gaps between high and low-income students and white and minority students were wide and showed few signs of improvement.[3] With hopes of initiating stronger growth in student performance, NCLB was signed into law on January 8, 2002 by the President with majority approval in the House and Senate.[4]

Under the NCLB accountability system, every state is required to set "challenging" academic achievement standards and annually test its students on core subjects.[5] In order to ensure that schools are effectively raising student performance and closing achievement gaps, NCLB orders that "states must develop a system of sanctions and rewards to hold districts and schools accountable for improving academic achievement".[6] Whether or not a school is rewarded or sanctioned largely depends on its

---

[1] *House Document 107-34,* web.
[2] Fusarelli, Lance D, "The Potential Impact of the No Child Left Behind Act on Equity and Diversity in American Education", *Educational Policy*. Vol 18, No. 1 71-94 (2004): pp. 72, web.
[3] *House Document 107-34,* web.
[4] Hursh, David, "Assessing No Child Left Behind and the Rise of Neoliberal Education Policies", *American Education Research Journal,* Vol. 44 No. 3 (2007): pp. 493, web.
[5] Public Law 107-110, web.
[6] *House Document 107-34,* web.

ability to make adequate yearly progress (AYP), which is derived from a minimum proficiency level that schools must achieve on state administered exams.[7] These proficiency levels must also apply separately to all demographic subgroups of students, not just the general student body. For example, if the white subgroup in a given school does not meet state proficiency levels, the entire school would fail to make AYP. Furthermore, schools that continually miss proficiency targets face a series of sanctions that increase in severity for every year the school fails to make AYP.[8]

In 2007, the Center on Education Policy (CEP) claimed to have "conducted the most comprehensive study of trends in state test scores since NCLB took effect."[9] The results presented were generally encouraging, noting that student achievement levels had steadily risen since NCLB's enactment in 2002. Additionally, they were able to measure pre- and post-NCLB trends in states with the appropriate data, finding that the majority had experienced more rapid year-to-year test score gains in the post-NCLB era. Their report also observed that overall, achievement gaps seemed to be narrowing, with only one state offering evidence that achievement gaps had widened.[10]

In this paper I aim to help answer the question of whether or not NCLB has had a significant impact in boosting student performance and closing achievement gaps. While the CEP's results certainly provide positive feedback concerning the effectiveness of NCLB, states may need more than just a span of five years to fully adjust to the new

---

[7] Public Law 107-110, web.
[8] Public Law 107-110, web.
[9] Chudowsky, Naomi, Victor Chudowsky, and Nancy Kober, "Answering the Question That Matters Most: Has Student Achievement Increased Since No Child Left Behind?", *Center on Education Policy,* (2007): pp. 1-96, web.
[10] Ibid. (2007)

accountability system. Allowing for this adjustment, I assess NCLB's effectiveness in the years 2008-2011 using the state of Texas as a case study.

A History of Texas Accountability

Many scholars agree that the nature of NCLB was based on the school accountability system implemented by Texas in the early 1990s.[11] Similar to NCLB, students in Texas were taught a state-mandated curriculum for which they were annually tested. On a school-by-school basis, test results were reported for both the general student body and each demographic subgroup. Based on this data, schools were given one of four performance ratings: exemplary, recognized, academically acceptable, or academically unacceptable.[12] This changed school accountability in two important ways. First, the Texas public school system became much more transparent. Underperforming schools could no longer hide from the public eye as they had before. Secondly, with public test score data broken down into racial and economic subgroups, the reality of existing achievement gaps between white and minority students was further exposed. Support systems were put into place by setting maximum student teacher ratios, requiring professional development programs for teachers, and offering half-day prekindergarten programs for children living in poverty.[13] This accountability system was monumental in that "while the rest of the nation was just beginning to consider the consequences of the

[11] Nelson, Sarah W., Narla W. McGhee, Lionel Meno, and Charles L. Slater, "Fulfilling the Promise of Educational Accountability, *Phi Delta Kappan,* Vol. 88, No. 09 (2007): pp. 702-709, web.
[12] Texas Education Agency, "Accountability Manual", Pub. No. GE6 602 02, (1996), web.
[13] Nelson, Sarah W., Narla W. McGhee, Lionel Meno, and Charles L. Slater, "Fulfilling the Promise of Educational Accountability, *Phi Delta Kappan,* Vol. 88, No. 09 (2007): pp. 702-709, web.

achievement gap, Texas was in the midst of developing a form model that would systemically alter the inherently inequitable nature of its public schools."[14]

The success story of Texas education reform was assessed in a report compiled by Achieve, Inc. in 2002. Using the National Assessment of Educational Progress (NAEP), a nationwide exam testing student knowledge[15], the study compared student achievement trends between Texas to those of other states. While noted as among the highest achieving states in NAEP math scores, "the state registered a "first" for the national exam--- its black students outperformed white students in a handful of other states."[16] In other words, the black students in Texas recorded stronger test results than the white students of several other states. Texas' students were sustaining strong improvement on their respective state exams as well: 70% of third grade students in 1997 passed their state tests, and when the same students were tested again as eighth graders, 80% passed.

Texas' efforts to close achievement gaps also appeared to be making considerable progress throughout the 1990s. Despite rapid growth in its minority populations, black and Hispanic students were experiencing much stronger achievement gains in academic performance than their white counterparts.[17] It is notable that Texas managed to sustain these gains in student achievement without an accountability system that relied on

---

[14] Nelson, Sarah W., Narla W. McGhee, Lionel Meno, and Charles L. Slater, "Fulfilling the Promise of Educational Accountability, *Phi Delta Kappan,* Vol. 88, No. 09 (2007): pp. 703, web.
[15] As of March 16th, 2013, the National Center for Education Statistics provided this information regarding the NAEP on its website, http://nces.ed.gov/nationsreportcard/about/
[16] Achieve, Inc, "Aiming Higher: Meeting the Challenges of Education Reform in Texas", (2002): pp 9, web.
[17] Ibid. (2002): pp. 37, web.

incentives and sanctions.[18] The Achieve, Inc. review board insinuated that peer pressure was an effective way to improve proficiency without the use of sanctions.[19] However, the provisions of NCLB have systematically altered the way Texas must monitor its schools. As with the entire nation, educational reform in Texas has significantly changed since NCLB's enactment in 2002.

Texas in the No Child Left Behind Era

As previously mentioned, NCLB allows states a certain degree of flexibility in setting their AYP standards and accountability measures. In Texas, AYP is derived from students' reading and math proficiency levels[20] on the Texas Assessment of Knowledge and Skills (TAKS), as well as demonstrated adequacy in either graduation or attendance rates. Additionally, Texas' NCLB requirements allow a "safe harbor" if schools fail the AYP threshold. "Safe harbor" is essentially a performance improvement indicator that allows for schools to pass AYP if the percentage of students scoring "not-proficient" in the previous year improves by 10%. This primarily helps schools at the bottom of the academic achievement distribution that would otherwise feel relatively hopeless in meeting yearly proficiency targets needed for AYP.[21]

---

[18] Kane, Thomas J. and Douglas O. Staiger, "The Primise and Pitfalls of Using Imprecise School Accountability Measures", *Journal of Economic Perspectives,* Vol. 16 No. 4 (2002): pp. 94, web.
[19] Achieve, Inc, "Aiming Higher: Meeting the Challenges of Education Reform in Texas", (2002): pp 9, web.
[19] Ibid. (2002): pp. 25
[20] *Note:* In the time period spanning 2008-2011, the years of focus in my study, TAKS was used. In 2012, Texas begun using the State of Texas Assessments of Academic Readiness (STARR) in combination with TAKS for setting AYP requirements.
[21] As of April 3rd, 2013, the Texas Education Agency's AYP website had provided this information regarding AYP requirements, [http://www.tea.state.tx.us/ayp/]

It should also be noted that *only* the schools receiving Title I federal funding must adhere to NCLB provisions. Title I funding originated from the aforementioned Elementary and Secondary Education Act to provide "financial assistance to local educational agencies (LEAs) and schools with high numbers or high percentages of children from low-income families"[22] The amount of Title I funds a school receives is dependent on a "statutory formula" that takes regional poverty levels and costs of education into account.[23] In 2011, the National Center for Education Statistics reported that 79% of Texas public schools were considered Title I.[24] For simplicity's sake, I will refer to Title I schools simply as 'schools' and will refer to non-Title I schools as such when specification is needed.

Unlike during the pre-NCLB era, schools in Texas that repeatedly fail to make AYP are sanctioned. The severity of such sanctions depends on the number of years a school has consecutively failed AYP. Once a school has failed to make AYP in back-to-back years, it is officially identified as a school in need of improvement and put under "stage one" sanctions. When this happens, administrators, teachers, parents, the LEAs, and outside experts are required to create a Campus Improvement Plan (CIP) evaluating the reasons for the school's failure. CIPs outline the way schools should manage curriculum, funds, and human capital so that its students have the highest likelihood of meeting the state's academic achievement levels in subsequent years. Furthermore, stage

---

[22] *Note:* This information was taken from the U.S. Department of Education's website on March 3rd, 2013, [http://www2.ed.gov/programs/titleiparta/index.html]
[23] *Note:* This information was taken from the U.S. Department of Education's website on March 3rd, 2013, [http://www2.ed.gov/programs/titleiparta/index.html]
[24] *Note:* This information was provided by the New America Foundation's website on April 14th, 2013, [http://nces.ed.gov/nationsreportcard/states/]

one sanctions require the LEA "to provide all students enrolled in the school with the option to transfer to another public school served by the LEA" that is not identified for poor performance.[25] When a school fails to make AYP for three consecutive years it is put under "stage two" sanctions. Beyond adhering to all requirements faced in stage one, schools must now also provide supplemental education services to its students. The Texas Education Agency (TEA) describes these services as free supplemental academic enrichment services (SES) "designed to increase the academic achievement of eligible children on the state assessment and to assist them in attaining proficiency in meeting the state's academic achievement standards."[26]

"Stage three" sanctions are incurred when a school has failed to make AYP for four years in a row. It is under these sanctions in which schools are exposed to some of the more intrusive components of NCLB. While still subjected to requirements enforced by stages one and two, the school is also identified as in need of "corrective action". Per stage three sanctions, the LEA gains significant control over school operations. With access to a higher level of management, the LEA attains the ability to fire school staff members it believes to be responsible for the failure to make AYP. Moreover, they also reserve the right to revise curriculum and appoint outside experts to assist teachers and administrators in raising student performance.[27]

---

[25] *Note:* Transportation to these schools must also be paid for by either the school itself or LEA. [http://www.tea.state.tx.us/ayp/]

[26] *Note:* The information in this paragraph was taken from the Texas Education Agency's AYP website on January 30th, 2013. [http://www.tea.state.tx.us/ayp/]

[27] *Note:* The information in this paragraph was taken from the Texas Education Agency's AYP website on January 30th, 2013. [http://www.tea.state.tx.us/ayp/]

When a school fails to make AYP for five consecutive years, "stage four" sanctions are subsequently enforced. Still subject to the previsions of previous sanctions, schools are additionally required to create a school-wide restructuring plan that would be fully prepared for implementation should the school fail in the sixth consecutive year. The TEA gives four primary examples of how a school could restructure itself; they can "(1) reopen the school as a charter school, (2) replace all or most of the school staff who are relevant to the failure to make AYP (which may include the principal), (3) contract for a private management company of demonstrated effectiveness to operate the public school, or (4) turn the operation of the school over to the State educational agency."[28] If the school in stage four fails to make AYP again the following year, it would be put under "stage five" sanctions and effectively implement the restructuring plan it had prepared.[29]

The severity and intrusiveness of sanctions was not intended to be punitive, but rather to be constructive. At its core, NCLB was enacted in order to help students at the bottom of the performance distribution that were not receiving sufficient academic attention. Using empirical analyses, I will evaluate NCLB's effectiveness based on its ability to stimulate achievement growth for low-performing schools in Texas. More specifically, I will empirically assess each sanction enforced by NCLB to better understand its impact in raising the overall proficiency level of students and closing achievement gaps. Going forward, I first analyze the important literature relevant to my

[28] *Note:* As of April 4th, 2013 this information regarding stage four sanctions was provided by the TEAs AYP section of their website. [http://www.tea.state.tx.us/ayp/]
[29] *Note:* The information in this paragraph was taken from the Texas Education Agency's AYP website. [http://www.tea.state.tx.us/ayp/]

study. Afterward I will explain the data used in my analyses as well as the methodologies implemented to perform them. Finally, I will thoroughly explain my results before concluding.

II. Literature Review

Prior literature has offered mixed results concerning the effects of federal sanctions on low performing schools and students. While the majority of studies demonstrate overall achievement gains due to the presence of an accountability system, undesired effects have been recorded as well. Randall Reback (2007)[30] studied schools' incentives to raise student performance levels when they were within range of achieving a higher accountability rating. Using data from Texas in the 1990s, he found that the lowest performing students made statistically significant improvements in test scores when their school had a reasonable chance of bettering its achievement rating. However, his analyses also suggested that the gains for students at the bottom of the performance distribution were influenced by reallocations of school resources. For example, if a school needed to boost its math performance, resources would be directed away from reading instruction and the needs of high performing students in order to raise the likelihood of low-performers meeting the required proficiency level. Brian Jacob (2005)[31] also used data predating NCLB by examining the effects of a school accountability policy implemented by the city of Chicago in 1996-1997. On the campus level, he discovered that low-performing schools responded better to the policy than their high-performing

[30] Reback, Randall, "Teaching to the rating: School accountability and the distribution of student achievement", *Journal of Public Economics,* Vol. 92 No. 5,6 (2007): pp. 1394-1415, web.
[31] Jacob, Brian A., "Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools", *Journal of Public Economics,* Vol. 89 (2005): pp. 761-796, web.

peers. Much like Reback, Jacob additionally found that it was students scoring at the 10th-50th percentiles that benefited the most from the accountability policy. While supporting Reback's findings, Jacob also provides evidence that resources were being directed away from the *very* lowest achievers with little hope to meet proficiency targets. Such findings have created concern over the "bubble kid" theory, suggesting that schools under accountability pressures will direct resources towards the underperforming students with the highest capacity to achieve proficiency levels.[32] However, a comprehensive investigation of the issue is beyond the scope of this paper.

Ladd and Lauren (2010)[33] observed similar results in the post-NCLB era using a North Carolina panel dataset with student, school, and yearly fixed effects.[34] Regression results found that low-performing students sustained significant gains in both reading and math scores after failing to meet AYP standards the prior year. In the same regression they detected adverse effects in high-performing students, suggesting that failure to make AYP spurred a reallocation of resources to low-achievers. Irrespective of whether or not this reallocation of resources is a matter of ethics,[35] the above research shows that NCLB has, at the *very* least, made progress in closing the achievement gap. I will use a similar empirical model to that of Ladd and Lauren, focusing instead on achievement gap trends between white and minority children.

---

[32] Booher-Jennings, Jennifer, "Below the Bubble: "Educational Triage" and the Texas Accountability System", *American Educational Research Journal*, Vol. 42, No. 2, pp. 231-268, web.
[33] Ladd, Helen F. and Douglas L. Lauren, "Status versus Growth: The Distributional Effects of School Accountability Policies", Journal of Policy Analysis and Management, Vol. 29, No. 3 (2010): pp. 426-450, web.
[34] *Note:* Given that I use this type of model in my analysis, its advantages will be later explained in the methodology section.
[35] Booher-Jennings, Jennifer, "Below the Bubble: "Educational Triage" and the Texas Accountability System", *American Educational Research Journal*, Vol. 42, No. 2, pp. 244, web.

In addition, I hope to offer renewed discussion of Hanushek and Raymond's study (2005)[36], which examined the effect of pre-NCLB accountability systems on white and minority achievement gaps. While their findings suggested that states with accountability systems recorded statistically significant gains in student achievement levels over states without accountability policies, their results hinted at a correlation between accountability systems and widening achievement gaps. Apparently, the effects varied "by subgroup, with Hispanics proportionally gaining most and Blacks gaining least. Because Whites gain more than Blacks do after accountability is introduced in absolute terms, the racial achievement gap actually widens with the introduction of accountability."[37] These results certainly add valuable insight to the literature, and beg the question of whether or not this trend has continued under NCLB. My research will attempt to understand such achievement gaps in a more current timeframe.

Relevant studies have also been performed by Gill, et al. (2009),[38] who used a regression discontinuity design (RD)[39] in assessing the impact of the NCLB accountability system in two unspecified states. At the state level, sanctioned schools did not produce statistically significant gains in student performance. However, they also tested the *threat* of being sanctioned; namely the response of schools that had only

---

[36] Hanushek, Eric A. and Margaret E. Raymond. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management*, Vol. 24, No. 2 (2005): pp. 297-327, web.
[37] Hanushek, Eric A. and Margaret E. Raymond. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management*, Vol. 24, No. 2 (2005): pp. 321, web.
[38] Gill, Brian, J.R. Lockwood III, Francisco Martorell, Claude Messan Setodji, Kevin Booker, "An Exploratory Analysis of Adequate Yearly Progress, Identification for Improvement, and Student Achievement in Two States and Three Citites", *National Longitudinal Study of No Child Left Behind,* "Technical Report" (2009): pp. iii-43, web.
[39] *Note: Gill*, et al. (2009) explain that their RD analyses "examine the relationship between the minimum AYP score for which a school is accountable and the school's subsequent achievement, assessing whether schools with minimum scores below AYP cutoffs (the treatment group) show achievement bumps in the subsequent year that distinguish them from the schools with minimum scores that exceed AYP cutoffs (the comparison group)".

missed AYP targets once.[40] They found that the students of these schools experienced statistically significant gains in proficiency for the previous year's lowest-scoring subgroup. Hanley Chiang (2007) also implemented a RD model to observe schools' responses to sanction threats in Florida. He too reported that sanction threats tended to have a statistically significant positive impact on a school's proficiency levels in the following year. However, using elementary reading as an example, he noted that improvements experienced from attending a threatened school "diminishes by more than half in the 2 years after the initial year of exposure."[41]

Fruehwirth and Traczynski (2012)[42] employ their RD design to instead study schools' response to the sanctions themselves, rather than the threat. Using a North Carolina dataset, they produce comparable results to those of Chiang. This study observed that after the first year of failure, schools experience statistically significant gains in test scores the following year. However, their analysis found that with repeated failure, schools were unable to maintain the same growth in proficiency gains that were experienced following the first year of failure. Interestingly, my analysis of Texas produces comparable results to that of Fruehwirth et al. (2012) and Chiang (2007). However, given that my empirical model is not constrained by the same sample size limitations in the above research,[43] I will additionally observe the effects of corrective action and restructuring sanctions on failing schools.

---

[40] *Note:* Recall that schools are only sanctioned after failing to make AYP for two consecutive years.
[41] Chiang, Hanley, "How accountability pressure on failing schools affects student achievement", *Journal of Public Economics,* Vol. 93 (2009): pp. 1045-1057,
[42] Fruehwirth, Jane Cooley and Jeffrey Traczynski, "Spare the Rod? The Dynamic Effects of No Child Left Behind on Failing Schools", *University of Wisconsin,* (2012): pp. 1-35, web.
[43] Ibid. (2009): pp. 14, web.

Application of a fixed effects model was used in a study by Zimmer et al. (2007)[44] to examine achievement gains in students from failing schools that had taken advantage of educational resources made available by NCLB sanctions; namely the option to transfer schools or utilize free supplemental education services (SES)[45]. Results from the analysis suggested that in the first year, supplemental educational services had a statistically significant positive impact on student test scores in both reading and math. Furthermore, students tended to experience even higher yearly rates of improvement on their exam scores in subsequent years. Achievement gains were felt by all racial and economic subgroups that exercised this option. On the contrary, significant gains in test scores were not detected for students who chose to transfer to better performing schools. However, the study noted that the sample size in this analysis was considerably smaller, and that its apparent statistical insignificance should be interpreted with caution. Heinrich et al. (2010)[46] similarly examined the effects of SES on student achievement gains using student-level fixed effects. Their results suggested that SES did not have a statistically significant impact on raising student performance, but it should be noted that the study was conducted on a different geographic location than that of Zimmer et al. (2007).

This paper will seek to understand the effectiveness of NCLB on raising student achievement in schools that continually fail. Given that the government spends upwards

---

[44] Zimmer, Ron, Brian Gill, Paula Razquin, Kevin Booker, J.R. Lockwood, "State and Local Implementation of the *No Child Left Behind Act*: Volume I- Title I School Choice, Supplemental Educational Services, and Student Achievement", *U.S. Department of Education* (2007): pp. iii-44, web.

[45] *Note:* In Texas, these resources are available to students attending schools that have failed to meet AYP for three consecutive years.

[46] Heinrich, Carolyn J., Robert H. Meyer, and Greg Whitten, "Supplemental Education Services Under No Child Left Behind: Who Signs up, and What Do They Gain?" *Educational Evaluation and Policy Analysis*, Vol. 32 No. 2 (2010): pp. 273-298, web.

of 14 billion dollars per year to fund NCLB[47], the economically important consequences of the new accountability system have stimulated extensive debate between policymakers and empirical researchers alike. Nonetheless, the degree of NCLB's success remains largely undecided. Drawing from previous research, I will give new insight on the widespread effects of NCLB on student performance and achievement gaps using a sanction-by-sanction methodology. Furthermore, this paper will evaluate the effectiveness of NCLB's most invasive sanctions in a way that, to my best knowledge, has yet to be discussed in the literature.

III. Data

In order to perform my analysis I used panel data from Texas in the years 2008-2011. School-level data for over 8000 public schools was collected from the Texas Education Agency's (TEA) AYP website. For each campus I have the percentage of students that met the AYP achievement standard on state examinations in reading and math for both the entire student body and relevant demographic subgroups. For schools that have failed to make AYP, I have data specifying which stage of sanctions the school was in. Additionally, my dataset contains student demographic information for individual schools as well as school-specific financial data that was provided by the TEAs Academic Excellence Indication System Reports (AEIS).

Prior to beginning my analysis, I generated several new variables. Firstly, I created binary variables to differentiate between Title I and non-Title I schools.[48] I also used binary variables to classify the stage of sanctions a school had been subjected to

[47] *Note:* This information was provided by the New America Foundation's website on April 14th, 2013.
[48] *Note:* Title I schools took the value of one, while non-Title I schools took the value of zero.

throughout the year. For example, the *stage3* binary variable would take the value of one if the school had been subjected to stage three sanctions. If the school had not been subjected to stage three sanctions, the variable would equal zero. In order to match the effects of these sanctions to the dependent variable in the *subsequent* year, I applied a one-year lag to the sanction-specific binary variables I had produced. Furthermore, I generated interaction binaries for Title I schools that were under sanction so that they could be assessed independent of non-Title I schools. I also needed to create a medium in which achievement gap trends could be assessed. To do this, I calculated the spread between the percentage of white students that had met the state performance standard and the percentage of black and Hispanic students that had met the state performance standard. This yielded four new variables, each one measuring the achievement gap between white and the aforementioned minority students in both reading and math proficiency.

For all schools in the dataset, the student demographic breakdown was approximately 36% white, 14% black, 46% Hispanic. On average, a school's student body had 87 and 81 percent of students meet the AYP threshold in reading and math, respectively. In my panel that spans 2008-2011, 69% of schools received Title I federal funding. Student demographics and performance levels were substantially different when distinguished between Title I and non-Title I schools. For a non-Title I school's general student body, 91% of students met the performance standard in reading while 83% met it in math. Demographically, non-Title I schools were 51% white, 12% black, and 31%

Hispanic with 37% of students being considered economically disadvantaged.[49] Additionally, non-Title I school tended to have approximately $10,555 in funds per pupil.

Remembering that Title I schools generally serve a higher percentage of low-income and minority students, their student achievement and demographic characteristics should not be surprising. Roughly 86% of a Title I school's study body satisfied performance requirements in reading and 80% of them did so in math. Students of Title I schools were typically 30% white, 14% black, and 54% Hispanic with 70% of students classified as economically disadvantaged. As compared to non-Title I schools' $10,555 in available funds per pupil, Title I schools generally had $6,193 per pupil.

Amongst all Title I schools in the dataset, just over 5% of them had been put under one of the five stages of federal sanctions. Of these sanctioned schools, students tended to be 11% white, 25% black, and 62% Hispanic while 79% of them were considered economically disadvantaged. Their average percentage of students meeting the required performance standard was 76% for reading and 56% for math. As a reference, Title I schools there were *not* under sanctions had 86% of students meet the performance standard for reading and 81% do so for math. These discrepancies are considerably large, which certainly justifies NCLB's motivation for sanctioning these low-performing schools. Finally, sanctioned Title I schools had about $6,385 in funds per pupil while Title I schools not under sanction had approximately $6,182, suggesting that sanctioned schools would receive additional money to help finance many of the additional campus-improvement requirements enforced on failing schools.

[49] *Note:* TEA classifies these students as students eligible for free or reduced-price lunch.

IV. Methodology

To perform my analyses, I use an ordinary least squares (OLS) regression model with campus and yearly fixed effects. Without fixed effects, it would be impossible to control for all campus-level characteristics that may influence student achievement. Indeed, omitted variable biases are often a primary reason for concern in a large panel data set. However, by using a fixed effects model, my regressions can control for unspecified covariates at the campus-level as long as they remain relatively constant throughout time.[50] I implement this methodology by first examining the effects federal sanctions have on student performance trends. The regression models for this analysis and their specifications are presented below[51]:

$$READPERCENTMET_{it} = \beta_0 + \beta_1 title1status_{it} + \beta_2 stage1\_lag_{it-1} + \beta_3 stage2\_lag_{it-1} + \beta_4 stage3\_lag_{it-1} + \beta_5 stage4\_lag_{it-1} + \beta_6 stage5\_lag_{it-1} + \beta_7 pctblk_{it} + \beta_8 pctwht_{it} + \beta_9 pcthisp_{it} + \beta_{10} pctecondis_{it} + \beta_{11} fundsperstudent_{it} + \delta_{kt} + u_{it} \quad \text{(I-V)}$$

$$MATHPERCENTMET_{it} = \beta_0 + \beta_1 title1status_{it} + \beta_2 stage1\_lag_{it-1} + \beta_3 stage2\_lag_{it-1} + \beta_4 stage3\_lag_{it-1} + \beta_5 stage4\_lag_{it-1} + \beta_6 stage5\_lag_{it-1} + \beta_7 pctblk_{it} + \beta_8 pctwht_{it} + \beta_9 pcthisp_{it} + \beta_{10} pctecondis_{it} + \beta_{11} fundsperstudent_{it} + \delta_{kt} + u_{it} \quad \text{(VI-X)}$$

in which

- $READPERCENTMET_{it}$ and $MATHPERCENTMET_{it}$ represent the percentage of students that meet AYP performance requirements for either a school's general student body, white subgroup, black subgroup, Hispanic subgroup, or economically disadvantage subgroup in the t^{th} year for both reading and math, respectively
- $title1status_{ij}$ differentiates between Title I and non-Title I schools

[50] Allison, Paul D. *Fixed Effects Regression Methods for Longitudinal Data Using SAS*. Cary, N.C: SAS Institute, (2005): pp. 1-7, web.

[51] *Note:* I would like to give credit to Zimmer et al. (2007) for the way I have formatted much of my methodology section.

- *stage1_lag$_{it-1}$, stage2_lag$_{it-1}$, stage3_lag$_{it-1}$, stage4_lag$_{it-1}$, and stage5_lag$_{it-1}$* indicate that the school pertaining to the i^{th} subgroup was put under one of five possible sanctions after failing to make AYP in the *t-1* year
- *pctwht$_{it}$, pctblk$_{it}$,* and *pcthisp$_{it}$* control for student race
- *pctecondis$_{it}$ and fundsperstudent$_{it}$*[52] control for income
- δ_{kt} captures campus-specific fixed effects over time
- u_{it} is the error term

The correct way to interpret these regression results may not be immediately apparent to unfamiliar eyes, so I will provide a hypothetical example for how it should be done. First, assume that the *title1status* coefficient has a value of 1.24. This would indicate that, per year, Title I schools tended to gain 1.24 percentage points *more* in student proficiency levels than non-Title I schools. Given that sanctions only apply to Title I schools, the sanctioned schools record this gain in "percent-proficient" independent of the gains or losses further caused by the sanctions themselves. It is the coefficient connected to one of the five sanction variables that will explain the effect the sanctions alone had on student proficiency levels. Taking this into account, assume that the *stage2_lag* variable carries a coefficient of 1.86. This would mean that on average, stage two sanctions stimulated an additional 1.86 percentage point increase in "percent-proficient" for the schools that were exposed to them. Therefore, schools under stage two sanctions would have gained 1.86 percentage points in student proficiency levels over Title I schools that were not under sanctions. Furthermore, by summing the *title1status* and *stage2_lag* coefficients, we would find that sanctioned schools gained 3.1 percentage points in student proficiency levels over non-Title I schools. While this method may seem tedious, it is advantageous in that it allows me to explicitly quantify the response a given

[52]*Note:* In Texas, school districts generate a considerable amount of their revenue from local property taxes. Given that different tax brackets represent different levels of income, I believe my use *fundsperstudentit* as an income control variable is valid.

school's proficiency levels had to stage two sanctions. If I had excluded the *title1status* variable, any gains or losses in proficiency levels caused by a school's Title I classification would likely be absorbed by the sanction variables, effectively diminishing their explanatory power.

Transitioning now to my achievement gap focus, I additionally ran four more regressions. The independent variables of these regressions remained the same as the ones above. However, I instead used the achievement gap metric that was explained in the data section to serve as the independent variable. The regressions for this study are presented below:

$$readspread\_whtafram_{it} = \beta_0 + \beta_1 title1status_{it} + \beta_2 stage1\_lag_{it-1} + \beta_3 stage2\_lag_{it-1} + \beta_4 stage3\_lag_{it-1} + \beta_5 stage4\_lag_{it-1} + \beta_6 stage5\_lag_{it-1} + \beta_7 pctblk_{it} + \beta_8 pctwht_{it} + \beta_9 pcthisp_{it} + \beta_{10} pctecondis_{it} + \beta_{11} fundsperstudent_{it} + \delta_{kt} + u_{it} \quad \text{(XI)}$$

$$readspread\_whthisp_{it} = \beta_0 + \beta_1 title1status_{it} + \beta_2 stage1\_lag_{it-1} + \beta_3 stage2\_lag_{it-1} + \beta_4 stage3\_lag_{it-1} + \beta_5 stage4\_lag_{it-1} + \beta_6 stage5\_lag_{it-1} + \beta_7 pctblk_{it} + \beta_8 pctwht_{it} + \beta_9 pcthisp_{it} + \beta_{10} pctecondis_{it} + \beta_{11} fundsperstudent_{it} + \delta_{kt} + u_{it} \quad \text{(XII)}$$

$$mathspread\_whtafram_{it} = \beta_0 + \beta_1 title1status_{it} + \beta_2 stage1\_lag_{it-1} + \beta_3 stage2\_lag_{it-1} + \beta_4 stage3\_lag_{it-1} + \beta_5 stage4\_lag_{it-1} + \beta_6 stage5\_lag_{it-1} + \beta_7 pctblk_{it} + \beta_8 pctwht_{it} + \beta_9 pcthisp_{it} + \beta_{10} pctecondis_{it} + \beta_{11} fundsperstudent_{it} + \delta_{kt} + u_{it} \quad \text{(XIII)}$$

$$mathspread\_whthisp_{it} = \beta_0 + \beta_1 title1status_{it} + \beta_2 stage1\_lag_{it-1} + \beta_3 stage2\_lag_{it-1} + \beta_4 stage3\_lag_{it-1} + \beta_5 stage4\_lag_{it-1} + \beta_6 stage5\_lag_{it-1} + \beta_7 pctblk_{it} + \beta_8 pctwht_{it} + \beta_9 pcthisp_{it} + \beta_{10} pctecondis_{it} + \beta_{11} fundsperstudent_{it} + \delta_{kt} + u_{it} \quad \text{(XIV)}$$

As noted above, the independent variables do not change. The same methodology from the first phase of regressions will be used to interpret these results. However, positive coefficients will now suggest a widening of the achievement gap while negative coefficients will suggest the opposite.

V. Results

Much like the literature that was discussed earlier, my results do not offer a definitive answer as to whether or not NCLB has been effective on all fronts. Regardless, my findings *do* suggest that federal sanctions have had a significant impact in raising performance levels for low-performing students across all subgroups. While these sanctions have had a much stronger impact on math performance, students' reading performance responded positively to sanctions as well. Being classified as a Title I school did not have a significant effect on proficiency levels in reading; however, it did seem to boost math performance among a few student subgroups.

The covariates also yielded interesting results. A school's "percent-proficient" in reading was negatively impacted at the 5% by a growing percentage of either black students or economically disadvantaged students. For math performance, none of the subgroups had a statistically significant negative impact on the "percent-proficient"; however, a larger white demographic tended to strongly influenced proficiency gains.

In general, the only worrying results from my analysis were the achievement gap responses to federal sanctions. For math, the majority of the sanction variables indicated that sanctions had a statistically significant impact on *widening* the achievement gaps between white and minority students. On the other hand, being classified as a Title I school had a statistically significant impact on closing the achievement gap in math performance. Given that Title I is not a direct product of NCLB, these findings are certainly troubling; however, they should not be entirely surprising. As mentioned in the literature review, Hanushek and Raymond (2005) observed that since white students gain

more in absolute terms than their minority counterparts when exposed to an accountability system, achievement gaps can widen even if all subgroups experience positive gains in performance.[53] Achievement gap trends in reading performance were perhaps even more dismal, especially since the coefficients for *title1status* were also positive in addition to the sanction variables.

Moving forward, I will now individually assess the impact each stage of sanctions has on the aforementioned dependent variables. For improved efficiency and comprehension in this section, I must make a few specifications concerning my use of terminology. I will refer to schools subjected to sanctions as "sanction schools". Or more specifically, if a given school had been put under stage two sanctions in the *t-1* year then they will be referred to as "stage two schools". The same will apply to all stages of sanctions and the schools they correspond with. Furthermore, Title I schools that succeed in making AYP will be called "adequate Title I schools", while non-Title I schools will still be referred to as such. Also, when referring to additional gains made by "sanction schools" or Title I schools, it should be known that these gains are *annual* and not an aggregate of the entire panel. Full regression results can be seen in Tables 1-3.

"Stage One" Analysis

As explained earlier in this paper, stage one sanctions require schools to perform a comprehensive evaluation to help understand the reasons it failed to make AYP. Once the problems have been identified, the school will create a Campus Improvement Plan (CIP) and strictly adhere to its guidelines. Additionally, the school must offer its students the

[53] Hanushek, Eric A. and Margaret E. Raymond. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management*, Vol. 24, No. 2 (2005): pp. 297-327, web.

option to transfer to another, higher-performing school within the district. Overall, my results indicated that stage one sanctions had a statistically significant positive impact in raising proficiency levels for both reading and math. However, the achievement gap trends in stage one schools was not optimistic.

Looking first at reading performance, my findings suggest that a stage one school's general student body had statistically significant gains at the 1% over both Title I schools that made AYP as well as non-Title I schools. For stage one schools, only the black subgroup of students did not experience statistically significant improvements in "percent-proficient". The white students from stage one schools tended to experience the largest proficiency improvements, gaining approximately 1.8 percentage points in "percent-proficient" over adequate Title I schools as well as non-Title I schools[54]. Additionally, stage one schools had a statistically significant impact (at the 10%) on the widening the achievement gap between white and Hispanic students. While my results hinted that there may have been a correlation between stage one sanctions and a *narrowing* of the achievement gap between white and black students, the *stage1_lag* variable was statistically insignificant.

Gains in math performance per stage one sanctions were generally more encouraging. However, just as in reading, black students were the only subgroup that did not achieve statistically significant gains in proficiency levels. Hispanic students from stage one schools sustained the biggest improvements, gaining an additional 2 percentage points in "percent-proficient" over both adequate Title I schools and 2.4 percentage

[54] *Note:* The *title1status* variable was statistically insignificant and carried a coefficient of close to zero.

points over non-Title I schools. Fortunately, my results indicated that a Title I classification resulted in a about a 1 percentage point narrowing in achievement gaps for both reading and math. However, achievement gap movements influenced by stage one sanctions were statistically significant at the 1%, indicating that achievement gaps had widened between white and the observed minority students.

"Stage Two" Analysis

To reiterate, stage two sanctions require schools to offer free supplementary educational services (SES) to its students while still adhering to the sanctions that were incurred in stage one. Starting with reading performance, a stage two school's student body experienced statistically significant gains in proficiency levels at the 5%. However, unlike stage one schools, the only subgroup in stage two schools that recorded statistically significant gains (at the 1%) as compared to adequate Title I schools and non-Title I schools was the white subgroup. My data suggests that stage two sanctions influenced a widening of the achievement gap[55] between white and Hispanic students, while apparently having no impact on the achievement gap between white and black students.

As for math, stage two sanctions had a significantly significant positive impact (at the 1%) on the proficiency levels of a school's general student body. Furthermore, my results indicate that all subgroups experienced statistically significant gains in "percent-proficient" by attending a stage two school. The black subgroup of students benefited the most, gaining an added 3.5 percentage points in proficiency over adequate Title I schools

[55] *Note:* The gap widening by about 2 percentage points as compared to adequate Title I schools and 2.5 percentage points over non-Title I schools.

and 4.5 percentage points over non-Title I schools. The response of achievement gaps in math to stage two sanctions was almost identical to that of achievement gaps in reading; the only difference being that a stage three schools' classification as Title I effectively narrows its achievement gap by an additional percentage point as compared to being classified as non-Title I. .

It is very encouraging to see black students having the strongest advances in math performance under stage two sanctions, especially because they were the only subgroup that did not have statistically significant gains in math by attending a stage one school. However, the fact that only the white subgroup from stage two schools made gains in reading performance is concerning. Since SES are introduced under stage two sanctions, these results perhaps suggest that teachers have not yet found an effective way to teach minority students reading material.

"Stage Three" Analysis

Again, "Corrective action" measures are taken in stage three. Under these sanctions, the local education agencies (LEAs) gain significant control over the stage three school by attaining the ability to fire inadequate staff members, replace curriculum, or exercise a number of other "campus improvement" plans. The impact stage three sanctions had on schools is very discouraging. For both reading and math, the only subgroup to achieve statistically significant results under stage three sanctions was the white subgroup. Their achievement gains were impressive, most notably the additional 5.6 and 6.4 percentage point gain in math "percent-proficient" over adequate Title I schools and non-Title I schools (respectively). Similar to achievement gap trends from

earlier sanctions, stage three schools also tended to experience a widening of achievement gaps.

While the performance gains made by white students are certainly impressive, the stagnancy in minority performance levels should raise red flags. The inherent cause for this is difficult to understand; however, it is possible that the invasive nature of "corrective action" measures may need more than the span of one year to have substantial effects on student performance. In order to verify this, I look to stage four schools to observe their response.

"Stage Four" Analysis

When in stage four, a school must prepare one of the four main restructuring plans[56] that would be ready for implementation if they were to fail to make AYP again in the following year. Although stage four sanctions do not add additional "campus improvement requirements" beyond those of previous sanctions, schools in stage four may begin to experience benefits from the far-reaching "corrective action" measures sustained in stage three. My results support this notion, as the stagnant performance gains experienced by all minority subgroups in stage three did not carry over to stage four.

In reading proficiency, a stage four school realized statistically significant gains for both the white and Hispanic subgroups at the 1%. Although stage four sanctions were still most effective in raising the proficiency levels of white students, Hispanic students

[56] *Note:* As mentioned in the introduction, a school can restructure itself in several ways: it can "(1) reopen the school as a charter school, (2) replace all or most of the school staff who are relevant to the failure to make AYP (which may include the principal), (3) contract for a private management company of demonstrated effectiveness to operate the public school, or (4) turn the operation of the school over to the State educational agency." This information was provided by the TEAs AYP website.

from stage four schools managed to gain almost 4 additional percentage points over both adequate Title I and non-Title I schools.[57] The unfortunate achievement gap trends for reading continue in stage four, but only in the achievement gap between white and Hispanic students.

Stage four sanctions generated strong achievement gains across the board for math proficiency[58]. Although the white subgroup of stage four schools felt the strongest improvements in "percent-proficient", the subgroup with the lowest performance gains[59] still managed to experience an additional 4.5 percentage point increase in proficiency levels over adequate Title I and non-Title I schools. Continuing the trend seen in earlier stages of sanctions, the achievement gap analysis indicated that stage four sanctions were the cause of widening achievement gaps in math. Despite this, my analysis has suggested that stage four sanctions have generated the much larger gains in proficiency levels than the sanctions prior.

"Stage Five" Analysis

Undoubtedly the most ambitious requirement enforced by NCLB on failing schools, campuses in stage five must put into action their carefully constructed restructuring plan. At this point, it is likely that school staff and teaching methods are substantially different from what they were a few years prior. According to my analysis, stage five sanctions stimulate much larger gains in student performance than any of the aforementioned stages of federal sanctions.

---

[57] *Note:* The *title1status* variable was statistically insignificant and carried a coefficient of close to zero.
[58] *Note:* All subgroups recorded statistically significant gains in math "percent-proficient" at the 1%.
[59] *Note:* This was the economically disadvantaged subgroup.

For reading, all subgroups of students except the black subgroup experienced statistically significant improvements in "percent-proficient" at the 1%. Of these, the Hispanic subgroup received the most benefit from stage five sanctions by gaining an additional 7.8 percentage points over both adequate Title I schools and non-Title I schools.[60] The response in math performance to stage five sanctions was even more pronounced, and statistically significant at the 1% across all four observed subgroups. White students of stage five schools had the strongest advances in proficiency levels. However, even the achievement gains of the economically disadvantaged subgroup, the lowest performers in stage five schools, managed to experience an additional gain of 9 and 9.8 percentage points over adequate Title I and non-Title I schools, respectively. Additionally, despite the fact that stage five sanctions did not seem to have statistical impact on closing achievement gaps in both reading and math performance, they are the only set of sanctions that did not seem to widen them.

VI. Conclusion

When NCLB was first enacted in 2002, its goal was to have all American students reach a proficient level of academic achievement by the school year ending in 2014.[61] That deadline is just one year away, and there is still a vast population of students that are far from reaching what would be considered "proficient". While well intentioned, both my research and the research conducted in past literature suggest that NCLB has not successfully tackled all of the problems it initially sought to solve.

---

[60] *Note:* The *title1status* variable was statistically insignificant and carried a coefficient of close to zero.
[61] Dee, Thomas S. and Brian Jacob, "The Impact of No Child Left Behind on Student Achievement", *Journal of Policy Analysis and Management,* Vol. 30, No. 3 (2011): pp. 418, web.

First and foremost, NCLB was created in response to the wide achievement gaps between low-income and high-income and minority and white students. Unfortunately, my study of Texas indicates that at the bottom of the student achievement distribution, achievement gaps have effectively widened. Whether or not this is applicable to the nation as a whole is irrelevant; NCLB was created to boost performance for *all* students. If achievement gaps in Texas seem to be widening, as my study suggests, then NCLB has not been effective in all respects. However, this does not mean that NCLB has failed. Plenty of studies, as well as my own analysis, give convincing evidence that students have experienced improved gains in academic achievement since NCLB took effect.

There are thousands of failing schools in the United States; my paper elected to gauge NCLB's effectiveness by assessing its capacity to boost student performance in these failing schools. The results suggest that the sanctions imposed by the new accountability system have had a statistically significant positive impact on student proficiency gains in the state of Texas. However, truly substantial gains in academic achievement were not felt until the later, more severe stages of federal sanctions were employed. As the nation moves forward in educational reform, my paper hopes to encourage policymakers and researchers to better understand the more obscure causes of continual school failure. If strategies could be developed to identify certain trends that make a school more vulnerable to repeated failure, perhaps the gains in student performance experienced in stages four and five would be realized much earlier.

Although my analysis gives evidence that all students are making strong advances in student performance, white students are clearly having the strongest gains. While a

comprehensive understanding of this issue is beyond the scope of this paper, it invites further investigation of this occurrence and its possible consequences. In the state of Texas, NCLB has generated uneven gains across the primary student subgroups. As our nation's leaders move forward in educational reform, careful measures must be taken to ensure that minority and low-income students are not left behind.

REFERENCES

Achieve, Inc, "Aiming Higher: Meeting the Challenges of Education Reform in Texas", (2002): pp 1-61, [http://www.achieve.org/files/Aiming_Higher_Texas.pdf], Accessed March 2013.

Allison, Paul D. *Fixed Effects Regression Methods for Longitudinal Data Using SAS*. Cary, N.C: SAS Institute, (2005): pp. 1-7, [http://support.sas.com/publishing/pubcat/chaps/58348.pdf], Accessed April 2013.

Booher-Jennings, Jennifer, "Below the Bubble: "Educational Triage" and the Texas Accountability System", *American Educational Research Journal,* Vol. 42, No. 2, pp. 231-268, [http://aer.sagepub.com/content/42/2/231.full.pdf+html], Accessed March 2013.

Center on Education Policy, "Answering the Question That Matters Most: Has Student Achievement Increased Since No Child Left Behind?", (2007): pp. 1-96, [http://www.cep-dc.org], Accessed February 2013.

Chiang, Hanley, "How accountability pressure on failing schools affects student achievement", *Journal of Public Economics,* Vol. 93 (2009): pp. 1045-1057, [http://www.sciencedirect.com/science/article/pii/S0047272709000693], Accessed February 2013.

Dee, Thomas S. and Brian Jacob, "The Impact of No Child Left Behind on Student Achievement", *Journal of Policy Analysis and Management,* Vol. 30, No. 3 (2011): pp. 418-446, [http://www.nber.org/papers/w15531], Accessed April 2013.

Fruehwirth, Jane Cooley and Jeffrey Traczynski, "Spare the Rod? The Dynamic Effects of No Child Left Behind on Failing Schools", *University of Wisconsin,* (2012): pp. 1-35, [http://www.econ.cam.ac.uk/faculty/fruehwirth/pubs/nclb_final_2012.pdf], Accessed February 2013.

Fusarelli, Lance D, "The Potential Impact of the No Child Left Behind Act on Equity and Diversity in American Education", *Educational Policy,* vol. 18, no. 1 (2004): pp. 72, [http://epx.sagepub.com/content/18/1/71.full.pdf+html], Accessed March 2013

Gill, Brian, J.R. Lockwood III, Francisco Martorell, Claude Messan Setodji, Kevin Booker, "An Exploratory Analysis of Adequate Yearly Progress, Identification for Improvement, and Student Achievement in Two States and Three Citites", *U.S. Department of Education* (2009): pp. iii-43, [http://www2.ed.gov/rschstat/eval/disadv/rd-ayp/report.pdf], Accessed February 2013.

Hanushek, Eric A. and Margaret E. Raymond. "Does School Accountability Lead to Improved Student Performance?" *Journal of Policy Analysis and Management*, Vol. 24, No. 2 (2005): pp. 297-327, [http://onlinelibrary.wiley.com/doi/10.1002/pam.20091/pdf], Accessed February 2013.

Heinrich, Carolyn J., Robert H. Meyer, and Greg Whitten, "Supplemental Education Services Under No Child Left Behind: Who Signs up, and What Do They Gain?" *Educational Evaluation and Policy Analysis*, Vol. 32 No. 2 (2010): pp. 273-298, [http://epa.sagepub.com/content/32/2/273.full.pdf+html], Accessed March 2013.

Hursh, David, "Assessing No Child Left Behind and the Rise of Neoliberal Education Policies", *American Education Research Journal,* vol. 44 no. 3 (2007): pp. 493, [http://aer.sagepub.com/content/44/3/493.full.pdf+html], Accessed March 2013

Jacob, Brian A., "Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools", *Journal of Public Economics,* vol. 89 (2005): pp. 761-796, [http://www.nber.org/papers/w8968], Accessed February 2013.

Kane, Thomas J. and Douglas O. Staiger, "The Primise and Pitfalls of Using Imprecise School Accountability Measures", *Journal of Economic Perspectives,* vol. 16 no. 4 (2002): pp. 91-114, [http://pubs.aeaweb.org/doi/pdfplus/10.1257/089533002320950993], Accessed April 2013.

Ladd, Helen F. and Douglas L. Lauren, "Status versus Growth: The Distributional Effects of School Accountability Policies", *Journal of Policy Analysis and Management,* Vol. 29, No. 3 (2010): pp. 426-450, [http://onlinelibrary.wiley.com/doi/10.1002/pam.20504/full], Accessed February 2013.

Nelson, Sarah W., Narla W. McGhee, Lionel Meno, and Charles L. Slater, "Fulfilling the Promise of Educational Accountability, *Phi Delta Kappan,* Vol. 88, No. 09 (2007): pp. 702-709, [http://www.pdkmembers.org/members_online/publications/Archive/pdf/k0705nel.pdf], Accessed March 2013.

Public Law 107-110, [http://www2.ed.gov/policy/elsec/leg/esea02/index.html], February 2013.

Reback, Randall, "Teaching to the rating: School accountability and the distribution of student achievement", *Journal of Public Economics,* vol. 92 (2007): pp. 1394-1415, [http://www.sciencedirect.com/science/article/pii/S0047272707000862], Accessed March 2013.

Texas Education Agency, "Accountability Manual", Pub. No. GE6 602 02, (1996), [http://www.tea.state.tx.us/index4.aspx?id=2147487432], Accessed March 2013.

U.S. House. 107th Congress, 1st Session. *House Document 107-34, A Report For Nationwide Education Reform Entitled, "No Child Left Behind"*. Washington: Government Printing Office, 2001.

Zimmer, Ron, Brian Gill, Paula Razquin, Kevin Booker, J.R. Lockwood, "State and Local Implementation of the *No Child Left Behind Act*: Volume I- Title I School Choice, Supplemental Educational Services, and Student Achievement", *U.S. Department of Education* (2007): pp. iii-44, [http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED497259&ERICExtSearch_SearchType_0=no&accno=ED497259], Accessed March 2013.

TABLE 1

Regression Results for Reading Performance	*Dependent variables*				
	(I)	(II)	(III)	(IV)	(V)
	Student body	Black subgroup	Hispanic subgroup	White subgroup	Economically disadvantaged subgroup
title1status	-0.04	0.70	0.01	0.07	0.13
	(0.24)	(0.53)	(0.35)	(0.28)	(0.31)
stage1_lag	1.19***	1.18	1.19***	1.76***	1.14***
	(0.32)	(0.76)	(0.45)	(0.48)	(0.41)
stage2_lag	0.96**	1.30	0.69	3.63***	0.83
	(0.48)	(1.13)	(0.66)	(0.75)	(0.61)
stage3_lag	-0.6	0.02	0.09	2.50**	-0.72
	(0.55)	(1.33)	(0.75)	(1.00)	(0.68)
stag4_lag	1.97***	0.002	3.84***	4.61***	1.44
	(0.78)	(1.72)	(1.08)	(1.44)	(0.98)
stage5_lag	4.47***	2.56	7.78***	6.39***	3.88***
	(1.04)	2.29	(1.42)	(1.93)	(1.30)
pctblk	-0.08***	-0.14**	-0.14***	0.001	-0.12***
	(.03)	(0.07)	(.04)	(0.04)	(0.04)
pctwht	0.04	-0.004	-0.04	-0.05	0.04
	(0.03)	(0.06)	(0.04)	(0.04)	(0.04)
pcthisp	-0.03	-0.03	-0.11***	-0.08**	-0.03
	(0.03)	(0.063)	(0.04)	(0.04)	(0.35)
pctecondis	-0.01**	-0.01	-0.001	-0.02**	0.008
	(0.01)	(0.02)	(0.01)	(0.01)	(0.008)
fundsperstudent	0.0001***	0.0000003	0.00002	-0.0002***	-0.00007*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	21279	12788	19514	16070	20759
R-squared (within)	0.01	.01	.01	.02	.01
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.10					

Table 2

Regression Results for Math Performance	*Dependent variables*				
	(VI)	(VII)	(VIII)	(IX)	(X)
	Student body	Black subgroup	Hispanic subgroup	White subgroup	Economically disadvantaged subgroup
title1status	0.50	1.08*	0.41	0.82**	0.84**
	(0.32)	(0.66)	(0.42)	(0.38)	(0.40)
stage1_lag	1.45***	1.48	2.00***	1.13*	1.39***
	(0.44)	(0.94)	(0.56)	(0.63)	(0.53)
stage2_lag	1.63***	3.34**	1.55*	2.16**	1.64**
	(0.65)	(1.42)	(0.82)	(0.98)	(0.79)
stage3_lag	1.11	1.92	1.25	5.64***	0.72
	(0.74)	(1.71)	(0.94)	(1.26)	(0.90)
stag4_lag	5.23***	6.68***	5.20***	9.02***	4.48***
	(1.05)	(2.18)	(1.32)	(1.78)	(1.27)
stage5_lag	10.45***	10.88***	9.46***	12.37***	9.16***
	(1.40)	(2.87)	1.75	(2.29)	(1.70)
pctblk	-0.05	-0.01	0.01	0.002	-0.08*
	(0.04)	(0.08)	(0.05)	(0.05)	(0.05)
pctwht	0.16***	0.18**	0.12**	-0.01	0.11**
	(0.04)	(0.08)	(0.05)	(0.05)	(0.04)
pcthisp	0.05	0.13*	0.07	0.02	0.06
	(0.04)	(0.08)	(0.05)	(0.05)	(0.04)
pctecondis	-0.002	-0.005	0.01	-0.02	0.02*
	(0.009)	(0.02)	(0.01)	(0.01)	(0.01)
fundsperstudent	0.0002***	0.00005	0.0002	-0.00008	0.0002***
	(0.00)	(0.00)	(0.00)	(0.0)	(0.0)
Observations	21354	13489	19889	16699	20975
R-squared (within)	0.13	0.06	0.12	0.05	0.14
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.10					

Table 3

Regression Results for Achievement Gap Trends	*Dependent variables*			
	(XI)	(XII)	(XIII)	(XIV)
	Reading Achievement Gap		Math Achievement Gap	
	White - Black	White - Hispanic	White - Black	White - Hispanic
title1status	0.84***	0.51***	-1.34***	-0.73***
	(0.29)	(0.20)	(0.35)	(0.24)
stage1_lag	-0.40	0.98*	2.14***	2.97***
	(0.70)	(0.52)	(0.84)	(0.63)
stage2_lag	-0.67	1.93***	0.09	1.84**
	(1.00)	(0.75)	(1.21)	(0.90)
stage3_lag	-1.78	1.17	4.34***	5.98***
	(1.39)	(0.97)	(1.68)	(1.17)
stag4_lag	0.61	2.79**	4.52**	4.02**
	(1.68)	(1.30)	(2.06)	(1.60)
stage5_lag	1.75	1.07	1.23	2.13
	(1.77)	(1.44)	(2.12)	(1.64)
pctblk	-0.03	-0.03	0.15***	0.06***
	(0.03)	(0.02)	(0.03)	(0.02)
pctwht	0.07***	0.01	0.13***	0.05**
	(0.02)	(0.02)	(0.03)	(0.02)
pcthisp	-0.002	-0.02	0.15***	0.05**
	(0.02)	(0.01)	(0.03)	(0.02)
pctecondis	0.02**	-0.01	-0.06***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
fundsperstudent	-0.0001**	0.00004	-0.0002***	-0.00007
	(0.00)	(0.00)	(0.00)	(0.0)
Observations	11943	16627	13136	17891
R-squared (within)	0.02	0.01	0.02	0.05
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.10				