MEASURING THE IMPACT OF SUPPLEMENTAL EDUCATIONAL SERVICES ON STUDENT ACHIEVEMENT

A Thesis

Presented to the faculty of the Department of Public Policy and Administration

California State University, Sacramento

Submitted in partial satisfaction of
the requirements for the degree of

MASTER OF PUBLIC POLICY AND ADMINISTRATION

by

Katharine A. Salling

SPRING
2012

MEASURING THE IMPACT OF SUPPLEMENTAL EDUCATIONAL SERVICES ON STUDENT ACHIEVEMENT

A Thesis

by

Katharine A. Salling

Approved by:

_____, Committee Chair
Su Jin Gatlin Jez, Ph.D.

_____, Second Reader
Robert W. Wassmer, Ph.D.

_____
Date

Student: Katharine A. Salling

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

_____, Department Chair _____
Robert W. Wassmer, Ph.D. Date

Department of Public Policy and Administration

Abstract

of

MEASURING THE IMPACT OF SUPPLEMENTAL EDUCATIONAL SERVICES ON STUDENT ACHIEVEMENT

by

Katharine A. Salling

This thesis examines the Supplemental Educational Services (SES) component of the *No Child Left Behind Act* of 2001 (NCLB) and focuses on the effect of this program improvement intervention on the academic performance of low-achieving students who participate compared to the academic performance of low-achieving students who are eligible but do not participate in the program.

Public schools and districts that serve a high percentage of students from low-income families receive federal funding under Title I of the *Elementary and Secondary Education Act* of 1965, reauthorized in 2001 as the *No Child Left Behind Act* (NCLB). Under NCLB, schools that receive Title I funds are required to test students annually to assess proficiency of basic skills in English-language arts and mathematics. Schools must demonstrate Adequate Yearly Progress (AYP) by increasing the margin of students who are academically proficient each year. Schools that fail to achieve the margin of proficiency after two years must reserve 15 percent of the Title I grant to provide Supplemental Educational Services (SES) for low-achieving, low-income students. The SES program is tutoring offered to students before- or after-school by private organizations not affiliated with the public school or district. Approved SES providers receive

payment for the tutoring services from the district's Title I funds that are reserved pursuant to the program improvement sanction.

NCLB requires the states to approve and authorize tutoring organizations or individuals to provide the SES program, and California requires firms to apply annually to be authorized and to submit accountability reports annually to maintain eligibility as a provider. In addition to information regarding providers' business operations and credential qualifications, SES providers must provide descriptive information regarding the tutoring services it delivers to individual students. However, the state or districts do not systematically evaluate SES providers to determine the statistical effect of the tutoring on students' academic achievement.

Whereas NCLB is a standards-based education reform policy, the absence of research-based evidence of the SES program effect creates ambiguity regarding the intent of the intervention strategy and provokes questions as to whether the policy meets the test of Pareto improvement. This thesis uses a linear regression analysis and administrator interviews to evaluate the estimated effect of SES participation on the academic achievement of students attending two program improvement schools in a small suburban school district located in Northern California.

The regression results indicate that a student's SES participation did not have a statistically significant effect on the student's test score for either English-language arts or mathematics. The conclusions drawn from the interviews with school district administrators indicate that school officials are uncertain and doubtful about the effectiveness of the SES intervention for improving student's academic performance on standardized test scores, particularly in comparison to program improvement strategies implemented in the classroom. With regard to implementation of the SES program, the interviews revealed that the district office assumes the role of ensuring

compliance with the NCLB mandates and the delivery of the SES program. The district conforms to the administrative recommendations and requirements of the state for managing the program but does not engage in additional monitoring or statistical evaluation of the effectiveness of the SES program or providers. School officials are generally satisfied with the way in which the district administers the SES program, yet they are hesitant to endorse the program as an effective use of resources in the absence of research-based evidence. Although the district does not conduct any formal survey to assess the level of satisfaction with the SES program among participants, the administrators report that informal communications with parents and students reveal generally positive opinions about the program.

I recommend a systematic approach for auditing the performance of the independent SES providers for evidence that the tutoring programs are effectively improving academic proficiency among the students who are eligible and participate in the entitlement program.

_____, Committee Chair
Su Jin Gatlin Jez, Ph.D.

_____
Date

ACKNOWLEDGEMENTS

I wish to thank the school district administrators and staff members who were so amenable and willing to assist me with gathering the data for this thesis. The prompt and courteous responses I received from each district and school employee were remarkable, and I am very grateful for the efforts made by each person to grant my requests and provide me with the information necessary to conduct this study. I also wish to thank the school board for their consent to access the district's data.

I wish to express gratitude to the Public Policy and Administration professors at California State University in Sacramento who provided exceptional instruction and advice to me while I completed the coursework and during the process of writing this thesis. I appreciate and am inspired by their dedication to the academic program and by their contributions to the public policy arena.

I wish to acknowledge and thank my husband and all of my family members for the understanding and patience they have shown me these past few years as I focused on achieving my goal of continuing my education. I am extremely grateful for their encouragement and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Tables Page

Chapter One

INTRODUCTION

State Academic Performance and Federal Proficiency Standards

Just as California public schools prepared to welcome students back to class for Fall 2011, State Superintendent of Public Instruction Tom Torlakson offered congratulations to a record-high number of schools that met or surpassed academic growth targets for 2010-11 based on the state's Academic Performance Index (API) (California Department of Education (CDE), 2011). Diffusing this good news, however, is the fact that an increasing number of California schools are failing to meet the federal accountability standard of Adequate Yearly Progress (AYP). California's API and the federal AYP are achievement benchmarks calculated from student test scores collected through the California Standardized Testing and Reporting Program. Each measurement takes a different approach to define achievement gains. The state's API establishes a school-level base score to track academic performance on assessment tests from one year to the next, targeting a minimum increase of five points per year for each school to reach a statewide goal of 800 to 1000 index points. By comparison, the federal AYP considers student-level progress by mandating that no less than 95 percent of students participate in standardized assessment testing each year and that at least two-thirds of students be proficient for 2011 in both English-language arts and mathematics (CDE, 2011). The particular challenge of AYP is that a school's percentage of proficient students must increase to reach 100 percent by the academic year of 2013-14.

Torlakson's announcement in August 2011 focused attention on the persistent problem that each year more of California's public schools fail to meet academic proficiency goals based on the accountability standard of AYP. This federal proficiency mandate is set forth in the *No Child Left Behind Act* (NCLB) of 2001, a reauthorization of the *Elementary and Secondary*

Education Act of 1965 (ESEA). Title I of the ESEA as reauthorized provides funding grants to schools that serve a high percentage of students from low-income families. As of 2011, approximately 62 percent of California's 9,858 public schools (6,157) receive Title I funds to support the educational needs of students disadvantaged by poverty. Yet, the majority of these schools (3,892 according to the CDE *Accountability Progress Report*, 2011) are not preparing the required percentage of students to achieve proficiency in English-language arts and mathematics.

The consequence for Title I-funded schools that do not exhibit adequate progress toward student proficiency after two consecutive years is mandatory program improvement (PI) intervention. Program improvement involves sanctions that require schools to set aside a total 20 percent of their Title I grant to provide transportation for students who choose to transfer to other schools and to provide students with the choice to enroll in private tutoring funded through the Supplemental Educational Services program (SES). Evidence is limited and uncertain as to the positive impact of the school choice and SES strategies to advance academic proficiency (Bathon & Spradlin, 2007), yet the NCLB policy continues to effect a growing number of schools and students.

In the current era of declining economies, there are increasing numbers of California's schools eligible to receive federal Title I funds to meet the education needs of disadvantaged students (from 6,066 in 2009 to 6,157 in 2011). And, as the number of Title I schools increases, there are also increasing numbers of these Title I schools remaining in and entering PI status for failure to achieve progress in student proficiency (from 2,796 PI schools in 2010 to 3,892 in 2011) (CDE, 2011). Thus, the NCLB requirement that PI schools must designate Title I grant funds for students to enroll voluntarily in private tutoring may have little or no impact on a school's progress toward increasing proficiency rates.

In this thesis, I focus on the SES mandate of NCLB and conduct a statistical analysis to find whether there is a measurable effect of the tutoring services on students' subsequent assessment test scores. The statistical findings will contribute to my evaluation of the SES provision as an appropriate use of Title I funds to promote proficiency among low-income students. In the remainder of this introductory chapter, I offer discussions relating to the intent of the NCLB policy in terms of equity, efficiency, and civil rights implications. I will describe the SES program and briefly discuss the regulatory policies adopted in California to ensure compliance with NCLB requirements. In concluding this chapter, I address the relevance and the imperative for quantifying the SES effect on assessment test scores as a method for evaluating the policy.

Providing Equal Access to Effective, High Quality Education

Public education reform is an important policy issue that consistently generates analysis and debate. Education policies, in general, involve mitigating the tension between politicians and practitioners and reconciling the conflict between values of equity and efficiency (Munger, 2000). A particular example is found in the federal *No Child Left Behind Act* (NCLB) of 2001, a reauthorization of the *Elementary and Secondary Education Act* of 1965. The political intent of the legislation is to promote improvement in academic performance and to eliminate achievement gaps among student subgroups and among schools that serve a high percentage of low-income families. The policy aims nobly to ensure equal access to quality public educational services for all students. The policy must meet standards of efficiency as well because it provides for appropriation of federal funds to state governments for allocation to local educational agencies serving communities disadvantaged by poverty. Arguably, NCLB program improvement interventions (requiring failing schools to divert their Title I funds to provide tutoring for low-achieving students, for example) are designed to create a Pareto improvement. That is, the

reallocation of Title I funds is expected to benefit low-performing students to a greater degree than the diversion of the funds results in a loss in benefit to low-income, academically proficient students. Ideally, low-achieving students will receive the extra instruction, improve academically, and bring the school's overall percentage level of proficiency up to meet the Adequate Yearly Progress standard.

However, although NCLB requires schools to provide education programs that are proven to be effective based upon state-developed assessment tests and other state accountability and monitoring systems (California's API, for example), the policy stops short of requiring rigorous analysis to measure the quality and effect of PI interventions, which would contribute definitively to the evaluation of Pareto efficiency. What is the significance of not measuring the SES effect? Since 2001, the ramification of NCLB is that the policy identifies and classifies an increasing number of schools serving low-income students as inadequate or failing based on the academic performance of a margin of students. If students' test scores are not reaching the cut point for proficiency following participation in federally funded supplemental tutoring, then the use of Title I funds for that purpose is not leading to Pareto improvement and is not, therefore, the best option for distributing educational resources targeted to assist disadvantaged students. The expansion of the SES program across Title I schools and districts each year creates a high-stakes need for evaluation of the SES effect in terms of assessing the rate of return in increased student proficiency.

Public education policies have civil rights implications (Burch, 2007), an additional aspect that commands attention to equity and efficiency. The implementation of policies that regulate the fair and equal distribution of public education dollars, and particularly policies to assist disadvantaged children, necessarily must include evaluative measures to control and prevent negative externalities stemming from decentralized financial interests that conflict with

the public benefit of education programs. Where the intent of NCLB is to promote and improve equal access to quality learning opportunities for all groups of students, from a broader perspective, consistently evaluating and improving public education will promote future economic development and advance the nation's ability to maintain an inclusive political democracy (Policy Analysis for California Education, 2008). Ironically, as an intervention for poor education outcomes among low-income students, NCLB sanctions such as SES further dilute the capacity for districts and schools to provide and improve their Title I educational programs. Not only are Title I funds diverted to pay outside SES providers, but there are also added administrative costs related to compliance and implementation of the provision which also impact districts and schools and may diminish their ability to offer, evaluate, or improve the quality of education programs.

Defining Supplemental Educational Services More Precisely

Public schools and districts throughout the nation that serve a high percentage of low-income students and receive NCLB Title I grant funds are required to evaluate students annually and report the percentage of students who demonstrate proficiency in reading and math. Title I schools that do not meet proficiency targets after two consecutive years must set aside 15 percent of Title I funds to provide their students access to academic tutoring, defined by NCLB as Supplemental Educational Services (SES). Students voluntarily receive SES outside of class time (before or after school) from a "public- or private-sector provider, including faith-based organizations, selected from a list of providers approved by the state" (U.S. Department of Education, 2004). NCLB prohibits program improvement schools and districts from providing SES directly to their students, unless the U.S. Department of Education approves a conditional waiver to do so.

NCLB requires that states must approve and authorize firms to provide SES. To comply, Title 5 of the California Code of Regulations requires local education agencies to "ensure" that a learning plan is developed for students enrolled in SES. Title 5 also vaguely requires SES providers to be accountable to the California Department of Education and "demonstrate a record of effectiveness in increasing the academic proficiency of students" (5 CCR §13075.4, 5 CA ADC §13075.4 and §13075.5). The regulation, however, does not explicitly require districts to verify or to report the effect of SES on individual student achievement in the same manner that schools must monitor progress in the classroom and report academic performance through standardized tests. In fact, according to a 2008 letter from the CDE to the USDE, Jack O'Connell and Theodore R. Mitchell wrote, "There is no practical methodology to separate the effect of the SES provider's contribution to academic proficiency from other factors" (O'Connell & Mitchell, 2008). Although the state annually updates its list of SES providers that are approved or unapproved based on provider applications, aside from state authorization, there is no rating system to evaluate the quality of SES providers. The absence of evidence on SES effectiveness creates significant uncertainty, particularly when schools remain in program improvement even after providing SES for students in successive years. Parents of students eligible for SES consequently have no information to compare SES provider agencies, which may influence participation rates down in that parents may not recognize the value of SES or, on the other hand, parents may opt for SES participation with unrealistic expectations about tutoring services based simply on a provider's approved status.

One aspect of California's efforts to comply with NCLB and to mitigate the sanctions related to the AYP-proficiency standard is the integration of the Academic Performance Index (API) as the state's additional indicator of AYP, required by NCLB. Established in California as part of the *Public Schools Accountability Act of 1999*, the API indicates a school's achievement

growth from the prior year's base score and thereby provides a means to compare schools. The API criteria, as employed for the federal AYP accountability reporting, applies to schools that educate and test a "numerically significant" subgroup of students that are at least 15 percent of the total students tested (more than 50 students of diverse ethnicities, English Learners, children with special needs, or those economically disadvantaged) (California Department of Education, 2011). As an additional indicator and based on standardized assessment tests, schools with diverse populations must achieve a growth API of at least 710 (for 2010-11) or must show growth in API of at least one point from the previous year. Thus, incorporating the API criteria for AYP allows a school to pass AYP through "safe harbor" if (1) the school has met the criteria of testing at least 95 percent of students in each subgroup, and if (2) the percentage of proficient students increased from the prior year by at least 10 percent in any of the student subgroups. In essence, a school that does not attain the AYP benchmark for total students exhibiting proficiency may still meet the API indicator for AYP if there is significant progress among subgroups. This model allows schools to recognize incremental improvements in proficiency among subgroups of students that are typically associated with academic achievement gaps, which ultimately contributes to a school's ability to identify which student groups could benefit the most from intervention strategies such as SES.

Relevance for Quantifying the SES Effect through Statistical Analysis

Evaluating the effect of SES on academic performance is a topic of national concern due to increasing numbers of PI schools required to provide SES (Bathon & Spradlin, 2007) and the increasing amount of limited Title I funds expended to independent firms to provide SES. U.S. Department of Education statistics indicate more than 3.6 million students were eligible to receive Title I SES as of 2006-07, which increased to more than 4 million in 2010. As for the tutoring industry, there were 3,050 state-approved SES providers nationwide in 2008, 88 percent of which

were private firms. With regard to expenditures, according to a survey of 300 districts and 1,483 schools nationwide, $375 million in federal funds were disbursed through Title I for SES during the 2005-06 school year (U.S. Department of Education: Office of Planning, Evaluation and Policy Development, 2009). With more schools each year remaining in program improvement status for failure to meet AYP, the evidence appears to indicate that these expenditures are not sufficiently improving academic proficiency among disadvantaged students.

The tensions that the NCLB mandates induce between politicians and education experts or other stakeholder groups raises the imperative to evaluate the policy's proficiency standards and obtain evidence of the SES effect using statistical analysis. President Barrack Obama and Education Secretary Arne Duncan announced in 2011 that states would be allowed to apply for discretionary conditional waivers to NCLB compliance beginning that academic year (2011-12), which revealed uncertainty at the highest levels about the success, equity, or efficiency of the policy's program improvement strategies (The White House, Office of the Press Secretary, 2011).

To inform the assumptions and objectives of this analysis, I report in Chapter 2 what we know about the SES effect according to academic research published since 2001, as the implications of NCLB have evolved. In Chapter 3, I describe the analytical methodology I use to evaluate the level of statistical significance of SES on students' standardized test scores. I explain the data I obtained to conduct this study, the construction of the regression model, and the variables included in the model and their predicted effects. In Chapter 4, I describe the results of the regression analysis and identify the magnitude and statistical significance of the variables in the model that influence student test scores. Concluding with Chapter 5, I discuss the findings of this case study and the implications of the research. I also address the limitations of the data and offer suggestions for expanding the analysis in the future.

Chapter Two

LITERATURE REVIEW

Research regarding the SES component of NCLB began soon after the act was authorized in 2001. Most of the earliest studies describe rapidly expanding demand for the intervention due to the increasing numbers of schools required to offer SES and students eligible for the intervention. Studies also described the expanding numbers of providers offering the tutoring services, as well as the implementation issues related to capacity, participation, and cost (Center on Education Policy, 2006; Government Accountability Office, 2006; Burch, Steinberg, & Donovan, 2007; Minnici & Bartley, 2007). Subsequent research has included empirical studies to evaluate the impact of SES on student academic achievement, which utilizes the expanding source of student-level data that state and local education agencies gather in compliance with NCLB assessment requirements.

To inform this thesis, I conducted a comprehensive review of the literature covering qualitative and quantitative research, which I will discuss in this chapter. First, I provide a brief chronological overview of the general findings related to the impact of SES on academic achievement according to statistical, quasi-experimental studies. I discuss the various methodologies developed and used by researchers to determine the extent of the SES effect. In the final two sections of this chapter, I summarize some of the main additional findings noted by researchers related to the SES program and review various recommendations researchers offer for future empirical strategies or analyses.

Chronological Overview of Research Findings

Researchers investigating the effect of SES since 2005 have revealed findings that range from no effect on academic achievement to a statistically significant but negligible effect. The studies that do report evidence of an effect indicate some association related either to SES content

(English-language arts or mathematics) or to the cumulative hours of participation in a tutoring program. In general, however, a comparison of the results of the research conducted during the past six years by third-party evaluators or by local public school districts reveals contradictory conclusions that warrant closer scrutiny of each report—for every analysis that finds SES has no significant effect on student achievement, there is another analysis reporting statistical evidence of a positive effect. In addition, the degree to which the research conclusions vary suggests that the perceived success or failure of the SES program is circumstantial relative to the time period of the data, based on the general assumption that the benefits of newly implemented policies may be delayed by challenges at inception that are resolved as the policy advances. The apparent inconsistencies noted upon reviewing all of the research conclusions also implies that benefits of the SES program are probably contingent upon factors present at the local level. Comparing the various findings regarding individual state or local education agencies with the findings from the meta-analyses and longitudinal studies substantiates these two points.

Examples of conflicting evidence about the effect of SES on academic achievement are seen in two studies that independently examined 2004-05 data from Minneapolis Public Schools and from Chicago Public Schools, respectively. Each study compared test scores of elementary school students who did or did not participate in SES, and the statistical analyses indicated that Minneapolis SES participants did not perform as well as non-participants (Heistad, 2005), while improvement in reading and math scores were found among the Chicago SES participating students (Jones, 2005). Two additional studies using 2004-05 education agency data, which examined SES programs for evidence of an association between academic improvement and specific providers, found zero or only limited significant effect. For example, there was no evidence of an effect in Louisiana for all but one SES provider, but the impact was found to be positive (Potter, Ross, Paek, Pribesh, & Nunery, 2006). The achievement data from Tennessee,

however, revealed no effect for all but two SES providers, but the statistics indicated a negative effect (Potter, Ross, Paek, McKay, Ashton, & Sanders, 2007). Another study, which examined data from an urban district in Kentucky with thirty program improvement schools required to provide SES in 2005-06, revealed the tutoring had no effect on year-end state test scores for reading or math (Munoz, Potter, & Ross, 2008). Similarly, a study exploring the SES effect on 2004-05, 2005-06, and 2006-07 test scores for Milwaukee Public Schools found no changes in academic performance associated with any amount of SES attendance compared with no SES attendance (Heinrich, Meyer, & Whitten, 2010). On the other hand, two consecutive studies conducted by the Los Angeles Unified School District using student data from 2004-05 and 2005-06 determined the SES program had a small but statistically significant effect on academic performance among participating students, with evidence that achievement gains were associated with cumulative SES attendance (Rickles & White, 2006; Rickles & Barnhart, 2007).

As referenced earlier, alongside the above-described independent studies of state and local education agencies offering SES is an interesting longitudinal meta-analysis conducted by the RAND Corporation for the U.S. Department of Education (USDE), which examined achievement data collected from nine large urban districts located throughout the nation for periods spanning 2000-01 through 2005. The research revealed that among five of seven districts with adequate data, students' reading and math scores improved after one year of participation in SES, with additional academic improvement gained after the second year of tutoring services (Zimmer, Gill, Razquin, Booker, & Lockwood III, 2007). Another longitudinal study commissioned by the USDE and released in January 2009 focused on two of five school districts, which had received waivers to the sanction prohibiting program improvement districts from acting as SES providers for their students. Located in Anchorage, Alaska, and Hillsborough County, Florida, these two districts provided SES as pilot programs beginning in 2006-07 to

expand the availability of tutoring for eligible students and to provide the USDE with information on the local effect of SES on academic achievement. Overall, SES participants in Anchorage did not exhibit statistically significant differences in academic achievement in reading or math when compared with eligible non-participating students. In Hillsborough County, small achievement gains were noted among SES participants in mathematics only, compared with eligible students who did not attend tutoring (Socias, deSousa, & LeFloch, 2009).

Later and subsequent analyses of the SES program offered to students attending Chicago Public Schools in 2007 and 2008 indicated that SES participation overall had no effect except among students who had tested below basic state proficiency standards in reading or math (Jones, 2009). Additional follow-up analysis of the SES program offered in Los Angeles in 2007-08 also revealed a continued small effect on student achievement in English-language arts and math across district populations of elementary, middle, and high school students eligible for SES (which represented 54 percent of the district population) (Barnhart, 2009). By comparison, results of a five–year longitudinal study using student-level panel data from one program improvement district, which was required to offer SES for students at five elementary schools and twelve middle schools between 2003-04 and 2007-08, revealed evidence of positive significant effects on math test scores among SES participating students but no significant effect on reading scores. The study also concluded a significant, cumulative SES effect on math and reading scores (Springer, Pepper, & Ghosh-Dastidar, 2009).

In summary, this section reviews and points out that the evidence and conclusions are mixed with regard to research results on the impact of SES to improve academic performance. In the next section, I will describe the methodologies used most commonly and in the most recent studies reviewed above and briefly discuss the advantages or limitations of each method as characterized by the respective researchers.

Review of Methods to Quantify the SES Impact

Each of the empirical studies previously described utilizes the statistical technique of linear regression to analyze the effect of SES and quantify the significance of the tutoring on student academic performance. In theory, a student's academic performance is dependent upon a combination of factors relating to the individual child and to their educational and social environments (Heinrich et al., 2009; Munoz et al., 2008). With linear regression, the researcher develops an equation to estimate mathematical coefficients for each variable that represents student factors (age, gender, aptitude, prior academic achievement), school factors (school type, teacher credentials, curriculum, academic or tutoring programs), and social factors (family structure or education level, ethnicity, language, income level). The equation holds all variables constant for all observations, and the coefficients then provide an estimate of the probable statistical influence or effect that each explanatory variable has upon the given outcome of academic performance.

In looking at the effect of the SES program, the key independent variable is SES participation, which is a categorical dichotomous or dummy variable in most studies, although some research includes a scale SES variable to estimate the impact of average SES participation by student-level hours of attendance. Logistic regression models are included in a few of the studies to estimate propensity scores for the predicted likelihood that students participate voluntarily in the SES program as a factor that may impact the outcome on the academic performance for both groups of students (participants and non-participants) (Heinrich, Meyer, & Whitten, 2010).

Researchers identify the dependent variable in each of the reviewed studies as a measure of academic performance based upon student-level standardized assessment test scores, separating math and reading scores in most studies, and often using the population mean and

standard deviation to calculate standardized scores (Barnhart, 2009). Each study includes test data for at least two school years, with the earlier year test score often included as an independent variable in the regression equation to control for the student factor of prior performance not related to SES participation. Regression analysis on the SES effect in Alaska and Florida considered longitudinal data over a period of five years (Socias, deSousa, & LeFloch, 2009); while another study examining the effect of attending particular SES programs combined test data compiled from four different years (Heinrich, Meyer, & Whitten, 2010).

Evaluating SES for the Los Angeles Unified School District, Rickles (2007) and Barnett (2009), controlled for student, school, and social variables (gender, language, special education status, grade level, SES program participation, meal program enrollment, parent education) and compared changes in California Standardized Test (CST) scale scores based upon three categories of eligible student groups (applied, attended, and completed SES). Using linear regression, the analysis included a calculation of the residual difference for each grade level between the expected and actual scores of participating and non-SES students; to improve the comparison across grade level and by student group, the researchers used overall mean and standard deviation to adjust and standardize scale scores for all students tested.

The 2008 study conducted for Tennessee Public Schools also used a "value-added" approach designed to control the student-ability effect and account for variances related to the school effect, thereby more precisely measuring the isolated effect of SES participation on achievement scores. In addition to student ethnicity and gender, the model also specifies grade level and matches SES participants against nonparticipants by comparable prior achievement to account for the probability that voluntary attendance has an effect on performance. A similar approach taken by Muñoz et al. (2008) for Kentucky Public Schools used 2005-06 data for a "quasi-experimental" model involving matched student pairs and pre and post-treatment test

scores. The Muñoz model also specifies ethnicity, prior achievement, and poverty based upon participation in the federal free lunch program. The Chicago Public Schools longitudinal five-year study (Jones et al., 2009) also evaluates the impact of SES using a value-added approach and describes benefits of Hierarchical Linear Modeling to quantify a combination of categorical and continuous variables. The equation design controls individual student, school, and social differences (gender, ethnicity, and grade level, performance on baseline achievement or other tests, disability, SES enrollment, discipline actions, attendance, district mobility, poverty status).

The regression analyses of the reviewed studies indicate a significance test to evaluate whether the coefficients for SES or other independent variables have an effect on standardized tests significantly different from zero, and a probability rate to determine the reliability that results reflect the true population. For example, four of the studies conclude that the effect of SES participation on post-treatment test scores was not significantly different from zero. However, the Los Angeles study indicated a small difference in adjusted test scores of eligible students who attended SES compared with eligible students who did not attend at the rate of 2 to 3 scale points. In discussing the absence of significant effect, Ross (2006) considers the possibility of low internal validity and power for detecting tutoring impacts as well as the inability to account for variance in student achievement attributed to academic environment (teachers, classrooms, schools). On the other hand, the Jones (2009) study indicates a small gain of 5 percent of the standard deviation associated with SES participation among the group of students with below standard performance in reading and math before the treatment. Barnett (2009) also found a small impact of SES participation at 5 percent of standard deviation. Ross (2008) addresses the problem of measuring the weighted scale effect of attendance, explaining that the threshold for mean differences in achievement of SES students compared with nonparticipants that is statistically meaningful is generally 20 percent of standard deviations or

greater in order to result in increased percentile scores that indicate levels of academic proficiency as required by NCLB.

Each analysis measuring the SES effect focuses some attention on the percentage of student participation or the number of hours students attend the tutoring program. Jones' study of Chicago schools considers SES attendance and dropout rates, offering insight regarding the demand for SES and the retention rate of SES providers. In Heinrich's 2009 evaluation of Milwaukee middle and high school students, increases in math and reading scores on standardized tests was associated with the number of hours students received SES from a particular vendor. With statistical significant at 4 percent of the standard deviation, however, the effect of tutoring hours is small in this study. For instance, the coefficient for the variable measuring attendance hours indicates a positive change in reading test scores by only .087 of a test unit for every additional hour of SES participation with a given provider.

Source and Limits of Data

As shown in Appendix A, the source of information for each study consists of large sets of student-level data gathered from state or local education agencies or SES providers. Most of the studies analyze summary statistics regarding student characteristics as well as SES enrollment, attendance, and program completion to qualify and explain findings and lack of statistical significance. Researchers also provide descriptive analyses for insight regarding variation in performance related to a student's choice to participate. Although the data represents all eligible students who may or may not participate in the program, the propensity to enroll, attend, and complete an SES program is associated to some extent with individual student characteristics as well as with student's school and social characteristics (Heinrich, 2009; Barnett, 2009; Jones, 2009). Excluding variables to measure propensity to attend, the frequency of SES attendance, or the mode of SES (one to one, small group, or web-based delivery of tutoring)

introduces bias and potentially limits the ability to draw causal conclusions. Aside from propensity, researchers also consider selection bias related to enrollment choice (also referred to as student discretion by Barnett, 2009), which disables the capacity for randomized field observations.

Recommendations from the Research Literature

As taken from the literature, recommendations to evaluate the impact of SES include designing a linear regression equation which specifies student performance in terms of variations in prior and post-treatment assessment test scores and which includes variables to describe student demographics including eligibility and participation in the SES program. Limitations noted in the research include the potential for selection bias because of the condition related to parents' choice to enroll or not to enroll their student in the SES program (and the extent to which that variable may ultimately influence academic performance more or less than SES attendance), which disables the capacity for randomized observations. As referenced earlier, methods recommended to reduce the potential selection bias error include using a quasi-experimental design that matches SES program participants with control students based on similar academic achievement and other student characteristics such as SES eligibility. Other suggestions to improve the regression equation involve better identification of SES provider characteristics and program modes to account for student propensity to enroll, attend, and complete the SES program (Jones, 2009; Heinrich, 2009). However, to the extent that SES providers might be considered partners in meeting the public policy objective of raising student academic achievement, the potential for using data produced by SES providers through pre- and post-treatment assessments is limited unless the SES programs, modes, and curriculum are standardized to allow for more accurate statistical evaluation.

Research measuring the effect of SES varies somewhat in terms of specific data source, method, or test statistic. Each successive study, however, contributes evidence to expand upon the current consensus that SES participation has a comparably minimal effect or may have a limited but cumulative effect for students with low proficiency who actively participate in consecutive years. In the following chapter, I will describe the analytical methodology I employ to measure the effect of the SES program implemented by the Northern California school district that provided the data for this study.

Chapter 3

METHODOLOGY

In this chapter, I describe the statistical method and data set I use to evaluate what impact that SES tutoring has on the standardized test scores of participating students enrolled at two program improvement schools in the same district in 2008 and 2009. I discuss the means by which I obtained the data to conduct this study, the theory for the construction of the regression model, the rationale for the variables included in the model, and my hypotheses about the effect of each variable within the given population. Later, I will discuss the results of the regression analysis in Chapter 4, wherein I also will identify the magnitude and statistical significance of the variables that may influence the student test scores within this data set.

Evaluating the SES Effect at the Local Level

As indicated by the literature, the emergent research on the effect of the SES program examines administrative or accountability data produced by state or local education agencies. As a case study, the subject school district is representative of California's local educational agencies in smaller suburban regions that receive Title I funding to serve low-income students. With two high schools, two middle schools, and eleven elementary schools, the district served more than 12,600 students in 2010-11. The student population consists of 50 percent White, 30 percent Latino, and nine percent African American children. District-wide, more than 36 percent of students participate in the free or reduced lunch program, 10 percent are English Learners, and 19 percent benefit from federal funding under Title I.

Table 1: District Enrollment, 2010-11 Academic Year

School Type	Number of Schools	Enrollment	Full-Time Teachers	Pupil-Teacher Ratio
Elementary	11	6,341	242.5	26.1
Middle	2	1,886	78.1	24.1
High School	2	3,753	143.5	26.2
K-12	1	465	24.2	19.2
Continuation	1	155	10.2	15.2
Community	1	20	2.0	10.0
Total	18	12,639	500.5	25.2

Source: California Department of Education

The U.S. Department of Education first identified the district for program improvement (PI) intervention under NCLB beginning in 2004-05 due to the failure of two elementary schools to meet AYP proficiency standards; the district was required to provide SES beginning in 2007-08. A student's SES eligibility is determined based upon student enrollment in a Title I school identified for program improvement and upon the student's California Standardized Test (CST) scores that fall in the range of basic or below basic proficiency. As required by NCLB, the district informs parents about eligibility for student enrollment in SES and provides parents with a list of the state-approved SES providers. To reiterate, the SES program is out-of-school tutoring intended as an academic intervention to increase the percentage of students who demonstrate academic proficiency in English-language arts and mathematics.

In 2007-08, the district set aside the mandated 15 percent of Title I funds to provide SES for eligible students. Following the 2008 standardized testing (which is the period covered in this study), the two elementary schools continued in PI status, and 136 of 851 eligible students enrolled in SES for the 2008-2009 school year. District records reveal that of the 136 students who enrolled in SES, there were 133 students who attended or completed one of thirteen private SES tutoring programs, receiving approximately 2,362 total hours of SES in aggregate at a cost of approximately $113,490, paid for from the Title I program funds set aside for the intervention.

The district continued in PI status in 2009-10 as the Title I schools again failed to meet AYP proficiency standards, and an additional elementary school identified for program improvement was required to offer SES. As a result, the number of SES eligible students increased to 893, with SES participation increasing to 186 students.

The district began the 2010-11 school year with all four of its Title I elementary schools in program improvement and with three required to set aside Title I funds to offer SES. Eligibility for SES tutoring, therefore, increased to 1,416 students. Despite the increasing number of eligible students, the SES participation rate in 2010-11 decreased to only 184 enrolled students, which is well below the enrollment threshold needed for the intervention to reach the percentage of students demonstrating proficiency according to AYP. Table 2 below summarizes the district's SES eligibility and participation rates.

Table 2: School District SES Program

Academic Year	Eligible Students	Participating Students
2008-2009	851	136
2009-2010	839	186
2010-2011	1,416	184

Source: District Records, 2011-2012

A statistical evaluation as to whether the SES program has a significant effect on participating students' test scores is potentially useful to the district in terms of compliance with NCLB monitoring requirements and in terms of contributing to the wider policy discussion regarding federal intervention to support academic progress. In addition to the statistical inferences, the data on student demographics, SES eligibility, and participation reveal the predominant student, school, and social conditions. The information may be useful to policymakers as well as administrators in terms of forecasting whether the SES intervention as implemented will be adequate in the future to improve aggregate academic proficiencies for the district's schools to meet AYP and progress out of PI status. The following section further

explains the methodology, data, and statistical model I employ to measure the effect of SES on student academic achievement.

Research Approach to Measure SES Impact

Following examples from the research literature, I apply the statistical technique of linear regression to measure the effect of the SES program on students' standardized test scores. Specifically, the sample for this study includes data pertaining to those students in grades 3 through 6 who were eligible to participate in the SES program given their attendance at either of two program improvement schools in 2007-08 and 2008-09 and their below basic proficiency according to performance on the 2008 California Standardized Test. Using student-level data from the school years cited above, my regression model takes a value-added approach by estimating a predicted 2009 test score based upon the 2008 test score and other explanatory variables held constant for SES eligible students. The difference between the predicted and actual 2009 test scores (also known as the residual) is the adjusted test score, or value added, which I use to estimate the effect of SES participation and to draw inferences about the impact of SES on academic proficiency.

Perceptions Regarding the SES Impact

In addition to the quantitative analysis, I report practical information learned during independent, one-on-one interviews with district administrators regarding their perceptions about the effectiveness of the intervention, service delivery and compliance with the program requirements, and participants' satisfaction with the program in general (Ross & Potter, 2006). With prior permission from district officials, I arranged to speak separately for approximately 30 to 45 minutes with each principal at the two program improvement schools required to offer SES in 2008-2009, the period of time and data used for this thesis (March 27, 2012). I also spoke for approximately 45 minutes with the district's chief officer of Education Services (April 2, 2012).

Several days prior to each meeting, I sent each person a consent form via email, which disclosed my topic of interest and reason for requesting the interview. I also provided each person with the interview questionnaire, which was identical for all three administrators (see appendix for consent form and questionnaire). At the beginning of each meeting, I advised the administrator that his or her identity would not be disclosed in this thesis report; each consented to participate in the interview and appeared to be interested in the topic as well as amenable and qualified to answer my questions to the best of his or her knowledge. I will discuss my findings and analysis from these meetings in Chapter 4 of this thesis.

The Statistical Model

Existing literature provides a theoretical basis for creating a regression equation using student-level data to evaluate the SES effect on assessment test scores. The dependent variable for this regression analysis is the California Standardized Test (CST) score for 2009 as reported for students in grades 3 through 6 who attended the district Title I schools required to offer SES tutoring in 2008 and 2009. Separate 2009 CST scores for English-language arts and mathematics provide the basis for assessing each student's academic achievement in these subjects. Achievement is also dependent upon, or a function of, a combination of explanatory variables such as the student's aptitude and past scholastic performance, the student's school environment, and the student's family economic and social inputs.

Individual 2008 CST scores are used in the equation to explain or account for how each student's aptitude and previous performance may impact his or her 2009 CST score. To account for the effect of school environment on student test scores, I include a variable to identify school enrollment and attendance as well as the key dichotomous variable of SES participation. Data related to individual student characteristics such as age (grade), gender, ethnicity, language, disability, participation in the free-lunch program, and level of parent education, are included to

account for the effect of individual characteristics, family economics and social inputs on academic achievement, which may confound the relationship between SES participation and CST scores (discussed in further detail in following sections). In applying the linear regression equation, the key independent variable of participation in the SES tutoring program will yield a coefficient. The coefficient estimates the statistical impact of SES on academic performance for this population and explains the variation in the students' 2009 CST scores, when holding all variables constant for each observed student and controlling for the other independent variables. In summary, the magnitude and direction of the SES coefficient will provide a means to evaluate whether SES had any impact on 2009 CST scores for the population of students in this data set.

Following is a discussion of the categorical variables I included in the linear regression equation, the assumptions justifying inclusion of the variables, and my hypothesis about the effect of each independent variable on the 2009 CST scores.

Student Aptitude and Past Scholastic Performance

To isolate the impact of scholastic aptitude on the 2009 CST, students' 2008 CST scores are included in the data to reflect past performance, which also serves as the basis by which the school district reports proficiency and considers student eligibility for SES participation. The 2008 CST score has an expected positive and significant impact on the 2009 CST score based upon the expected value of familiarity with the test structure as well as exposure to relevant standardized curriculum and test preparation activities in the classroom prior to 2009 testing.

School Environment

To explain variations in the 2009 CST scores related to learning environment, I have included a dichotomous variable to represent a school factor defined by enrollment/attendance at either of the district's two program improvement elementary schools required to provide SES in 2008 and 2009. These two schools served similar populations of students from comparable

socio-economic and ethnic neighborhoods. In 2008-09, they reportedly served a combined population of 1,332 students, with more than 60 percent of these students socio-economically disadvantaged, 57 percent Hispanic, and 39 percent English language learners proportionately distributed. An additional similarity is that in 2008-09, each school had a student-to-fully credentialed teacher ratio of approximately 17 to one. One elementary school served a majority of the Title I students, both in terms of total enrollment between the schools and in terms of the total observations for this study. The school enrollment variable is dichotomous and equal to one for students attending and tested at School Site I (76 percent of students in this data set). The school a student attends is likely to have a positive or negative effect on the academic performance. A negative effect may be more likely if the school receives federal Title I funds compared to a school that does not based upon the assumption that students attending Title I program improvement schools may not perform as well on the standardized tests due to challenges associated with low socio-economic factors and limited opportunities for individual academic or extra-curricular activities. I expect the statistical effect of school attended to be positive in this case, but the effect may not be significant given the similarities between the two school environments.

The key variable of SES participation is also dichotomous and defined as students who enrolled and attended at least one hour of tutoring during the period between the 2008 and the 2009 standardized tests. I include SES participation in the category of school environment based upon the assumptions that the tutoring experience represents an extension of students' academic activities and that the school administrators or faculty may encourage SES participation or have some influence on whether students volunteer or their parents chose to enroll them in SES tutoring. I also consider the SES variable part of the school environment based upon the possibility that administrators or faculty interact with SES providers directly or indirectly in an

effort to align tutoring activities with classroom or standardized curriculum. Based upon my review of the published studies on the effect of the SES program, I hypothesize that the factor of SES participation has an impact on 2009 CST in this population, which is probably minimal or insignificant.

Student Characteristics, Family Economic, and Social Inputs

The factors of individual student characteristics that may explain variation in the CST scores include gender, ethnicity, age (expressed by grade level), and disability status. As dichotomous or dummy variables in the regression equation, the outcome for male students (52 percent of the population), Hispanic students (88 percent), and grade 4 students (29 percent) establishes a point of reference for the alternative(s) of these variables. My hypothesis about the effect of these variables are that female students may score better than male students on the test for ELA but not math; Hispanic students may not perform as well as non-Hispanic due to possible language barriers; and younger students in the grades 3 or 4 may perform less well than older students. Of these variables, I hypothesize that age is likely to have more statistical effect on test scores than gender or ethnicity. With regard to disability status, my hypothesis is that disability status is likely to have a negative and significant effect on the 2009 standardized test scores.

Variation in 2009 CST scores is dependent upon individual student characteristics that are related to the student's economic and social environments as well. Examples of these factors include educational attainment of the student's parents; family income level; and language fluency of the student (related to language fluency of the student's family). A factor to measure student mobility (immigration to the subject school from another school or district) is also included to estimate how the student's social environment potentially disrupts academic activity and impacts performance on the standardized tests. To capture parent education attainment for each observation, I designated the value of one if a student's parent reported not graduating high

school, being a high school graduate, having college experience, or if the student's parent did not report education level. I hypothesize that a student whose parent graduated high school or has college experience is likely to score higher on the 2009 CST than a student whose parent did not report educational level, did not graduate high school, or go to college. Family income is also a dummy variable with the value of one if the student participates in the district's free-or-reduced lunch program (83 percent of this population), which I hypothesize will have a negative effect on those student's 2009 CST score. To specify the effect of language fluency on 2009 CST scores, the data includes a dichotomous value for each of the four fluency categories, with the category of students who speak English only indicated as the reference for comparison (52 percent). My hypothesis with regard to language is that a student who is Spanish-English bilingual or reclassified as English fluent will score better than a student who is learning but not yet English fluent.

The theoretical regression equation is presented below and includes my hypotheses about the expected positive (+) or negative (-) effect in terms of variation in 2009 CST scores. Where the hypothesis is that the variable will have an effect that is either positive or negative, the null hypothesis is that the respective variable will have an effect on test scores that is equal to zero. Likewise, where the hypothesis is that the variable will have a positive effect, the null hypothesis is that the variable will have an effect equal to or less than zero; and where the hypothesis is for an effect that is negative, the null hypothesis is that the effect will be equal to or greater than zero.

2009 CST Scores = f [Student Aptitude and Past Scholastic Performance + School Environment + Student Characteristics, Family Economic and Social Inputs]

Student Aptitude and Past Scholastic Performance = f [2008 CST Scores (+)]

School Environment = f [Title 1 Program Improvement School Enrollment (+),
SES Participation (+/-)]

Student Characteristics, Family Economic and Social Inputs = f [Gender (+/-),
Ethnicity (+/-), Disability (-), Grade (+/-), Parent
Education (+/-), Lunch Program Participation (-), English
Language Fluency (+/-), Mobility (+/-)]

Data Sample

The data set I used for this analysis includes the 2008 and 2009 student-level information collected and reported by the subject school district as required by the California Department of Education. Although the data initially received from the district included approximately 900 aggregated observations, many variables were missing for several individual observations. Matching identification numbers, I collected only the data which presented an observation of all variables for each student as previously described, which includes 2008 and 2009 CST scores, school of enrollment, SES participation, and other student, family, economic and social explanatory factors. The final sample size consists of 325 observations with test scores for two years in both English-language arts and mathematics. Of the 325 students with complete data for all relevant variables, only 61 students were SES participants. Table 3 below presents the variable labels for the data and a description of each variable used in the regression equation.

Table 3: Variable Labels, Descriptions, and Sources

Variable Label	Variable Description
Dependent Variables	Data Source: School District
2009 ELA CST Score	Student-level score on 2009 California Standards Test (English-Language Arts Content Assessment)
2009 Math CST Score	Student-level score on 2009 California Standards Test (Mathematics Content Assessment)
Independent Variables	Data Source: School District
Student Academic Aptitude:	*Data Set: Student-Specific Numeric Identification Data*
2008 ELA CST Score	Student-level score on 2008 California Standards Test (English-Language Arts Assessment)
2008 Math CST Score	Student-level score on 2008 California Standards Test (Mathematics Content Assessment)
School Environment:	*Data Set: School Specific Enrollment Data*
School Site A	Student Attended and Tested at School Site A in 2008 and 2009 = 1
School Site B	Student Attended and Tested at School Site B in 2008 and 2009 = 0
Supplemental Educational Services (SES) Participation	Key independent variable for participation in at least one hour of tutoring following the 2008 and prior to the 2009 CST = 1; Non-SES Participant = 0
Social Factors:	*Data Set: Student-Specific Identification Data*
Gender	Male reference = 1; Female = 0
Ethnicity	Hispanic reference = 1; Other = 0
Disability	Student with disability = 1; No disability reported = 0
Grade 3	Student Grade 3 at 2009 CST = 1; Other = 0
Grade 4 (Reference)	Student Grade 4 at 2009 CST = 1; Other = 0
Grade 5	Student Grade 5 at 2009 CST = 1; Other = 0
Grade 6	Student Grade 6 at 2009 CST = 1; Other = 0
Free/Reduced Lunch	Participant in Free or Reduced Lunch Program = 1; Not participating in lunch program = 0
Parent No High School	Parent reported not graduating from high school = 1; Other = 0
Parent High School Education	Parent reported graduating from high school =1; Other = 0
Parent Has College Experience	Parent reported attending college = 1; Other = 0
Parent Education Not Reported	Parent declined to report education level = 1; Other = 0
English Language Only	Student's language is English only in 2008 = 1; Other = 0
English Second Language	Student classified as Initially-Fluent English Proficient (English as a second language) in 2008 = 1; Other = 0
Reclassified Fluent-English Language	Student is reclassified from English Learner to Fluent-English Proficient in 2008 = 1; Other = 0

Table 3: Variable Labels, Descriptions, and Sources - Continued

Variable Label	Variable Description
English Learner	Student is English Learner not yet English Fluent Proficient = 1; Other = 0
School Mobility	Student attended another school within the district prior to 2008 = 1; Other = 2
District Mobility	Student attended another district prior to 2008 = 1; Other = 0

Following is Table 4, in which I present the descriptive statistics for the variables including the mean, standard deviation, and minimum/maximum value calculated for each. Later, in Chapter 4 I will provide a detailed analysis of the data and results of the linear regression, as well as a discussion of the information discovered through personal interviews with district administrators.

Table 4: Descriptive Statistics

Variable Label	Mean	Standard Deviation	Minimum Value	Maximum Value
2009 ELA CST Score	323.75	50.33	222	476
2009 Math CST Score	334.26	69.15	182	600
Student Academic Aptitude				
2008 ELA CST Score	316.74	48.24	202	525
2008 Math CST Score	342.52	73.82	178	600
School Environment				
School Site A	.2431	.4296	0	1
School Site B	.7569	.4296	0	1
SES Participation	.1877	.3911	0	1
Social Factors				
Gender (Male Reference)	.5185	.5004	0	1
Ethnicity (Hispanic Reference)	.8827	.3222	0	1
Disability Status	.0185	.1350	0	1
Grade 3	.2646	.4418	0	1
Grade 4 (Reference)	.2123	.4096	0	1
Grade 5	.2892	.4241	0	1
Grade 6	.2338	.4239	0	1

Table 4: Descriptive Statistics - Continued

Variable Label	Mean	Standard Deviation	Minimum Value	Maximum Value
Free/Reduced Lunch Program	.8277	.3782	0	1
Parent Not High School Graduate	.2369	.4258	0	1
Parent High School Graduate	.3846	.4873	0	1
Parent Has College Experience	.2677	.4434	0	1
Parent Education Not Reported (Ref.)	.1108	.3143	0	1
English Language Only (Reference)	.2462	.4314	0	1
English Second Language	.0532	.2230	0	1
Fluent-English Reclassified	.1838	.3879	0	1
English Learner	.5169	.5005	0	1
School Mobility	.4123	.4930	0	1
District Mobility	.2554	.4367	0	1

In using linear regression analysis to estimate the impact of the SES on student test scores, the Ordinary Least Squares technique requires that all independent variables be uncorrelated with the error term of the equation. The correlation coefficient (r) measures the strength and direction of the linear relationship between each of the explanatory variables to detect multicollinearity. The closer the absolute value of the coefficient is to 1, the stronger the correlation is between the two variables, which indicates that the two variables move statistically somewhat relative to one another either in the same or in opposite directions. In this study, with an equation comprised of multiple independent variables, I use the threshold of $r \geq .50$ to assess potential multicollinearity. Thus, when two explanatory variables have a correlation coefficient of greater than or equal to .50, there is potential multicollinearity between the variables that may inflate the standard errors of the coefficients and indicate the variables have no statistically significant effect on the outcome. Below are Tables 5 and 6, each presents a matrix of statistically significant correlation coefficients at the level of .05 and .01 for the English-Language Arts and mathematics regressions, respectively.

Of the independent variables used in the equation to estimate the impact of SES participation on the 2009 CST scores for English-language arts, there is potential negative correlation between the variables for 2008 CST score and English Language Learner (-.5374), which indicates that as one of these variables increases, the value of the other decreases. English Speaker is also potentially correlated negatively with the variable for Ethnicity (-.5256), English Learner (-.5911), and Reclassified Fluent in English (-.4922). There is a potential positive correlation found between the School Mobility and District Mobility (.6992), which are variables included to signify whether the student attended another school or district prior to the 2008 CST and to account for how disruptions in school attendance may impact test scores. Similarly, the variables used to estimate SES participation on 2009 CST scores for mathematics indicate potential negative correlation between English Speaker and Ethnicity (-.5256), English Learner (-.5911), and Reclassified Fluent in English (-.4922), as well as a positive correlation between School and District Mobility (.6992). In Chapter 4, I again address the issue and conduct statistical tests using STATA software to identify the existence of bias due to multicollinearity.

Table 5: Correlation Coefficients for Explanatory Variables Effecting 2009 CST for English-Language Arts

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	1	-.18*	-.19*		-.20*	-.12	-.11	-.13	.24*				.12	-.13	.23*	.20*	-.54*	.32*			
2	-.18*	1		-.12	.40*										-.26*		.23*			-.12	
3	-.19*		1			.23*	.23*								-.11		.18*		-.12	-.13	
4		-.12		1		.13										-.02					
5	-.20*	.40*			1										-.53*		.34*				
6	-.12		.23*	.13		1													-.41*	-.11	
7	-.11		.23*				1	-.31*	-.38*	-.33*							.19*	-.18*		-.29*	-.16*
8	-.13						-.31*	1	-.33*	-.29*											.11
9	.24*						-.38*	-.33*	1	-.35*							-.16*	.12			
10							-.33*	-.29*	-.35*	1								.11		.26*	
11											1										
12					.13	.13					-.44*	1			-.11						

Table 5: Correlation Coefficients for Explanatory Variables Effecting 2009 CST for English-Language Arts – Continued

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
13	.12	-.13	-.11		-.23*			.11			-.34*	-.48*	1								
14	-.13	.18*									-.20*	-.28*	-.21*	1							
15	.23*	-.26*	-.11		-.53*							-.11	.35*	-.18*	1						
16	.20			-.13											-.13	1					
17	-.54*	.23*	.18*		.34*		.19*		-.16*				-.19*		-.59*	-.24*	1	-.49*			
18	.32*						-.18*		.12	.11					-.27*	-.11	-.11	1			
19			-.12			-.41*													1		
20		-.12	-.13			-.12	-.30*			.26*					.12					1	
21							-.16*	.11							.12					.70*	1

Correlation Coefficients listed are significant at the .05 level unless denoted with an asterisk (*) for the .01 level. *Legend*: 1=ELA 2008 CST Score, 2=School Site Enrollment, 3=SES Participation, 4=Gender, 5=Ethnicity, 6=Disability, 7=Grade 3, 8=Grade 4, 9=Grade 5, 10=Grade 6, 11=Parent Not High School Grad, 12=Parent High School Grad, 13=Parent College Educated, 14=Decline Educ. Report, 15=Student English Speaker, 16=Student Spanish-English Speaker, 17=Student English Learner, 18=Student Reclassified Fluent Learner, 19=Student Free/Reduced Lunch Participant, 20=School Mobility, 21=District Mobility

Table 6: Correlation Coefficients for Explanatory Variables Effecting 2009 CST for Mathematics

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	1	-.13	-.18*		-.17*			-.15*		-.21*			.16*		.16*	.13	-.41*	.27*			
2	-.13	1		-.12	.40*								-.13	.17*	-.26*		.23*			-.12	
3	-.18*		1			.23*	.23*						.11		-.11		.18*		-.12	-.13	
4		-.12		1		.13										-.13					
5	-.17*	.40*			1							.13	-.23*		-.53*		.34*				
6			.23*	.13		1						.13							-.41*	-.11	
7			.23*				1	-.31*	-.38*	-.33*							.19*	-.18*		-.29*	-.16*
8	.15*						-.31*	1	-.33*	-.29*			.11								.11
9							-.38*	-.33*	1	-.35*							-.16*	.12			
10	-.21*						-.33*	-.29*	-.35*	1								.11		.26*	
11											1	-.44*	-.34*								
12					.13	.13					-.44*	1	-.48*		-.11						

Table 6: Correlation Coefficients for Explanatory Variables Effecting 2009 CST for Mathematics – Continued

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
13	.16*	-.13	-.11		-.23*			.11			-.34*	-.48*	1		.35*		-.19*				
14		.18*									-.20*	-.28*	-.21*	1	-.18*						
15	.16*	-.26*	-.11		-.53*							-.11	.35*	-.18*	1	-.13	-.59*			.12	
16	.13			-.13											-.13	1	-.24*				
17	-.41*	.23*	.18*		.34*		.19*		-.16*				-.19*		-.59*	-.24*	1	-.49*		-.13	
18	.27*						-.18*		.12	.11					-.27*	-.11	-.49*	1			
19			-.12			-.41*													1		
20		-.12	-.13			-.12	-.30*			.26*					.12		-.13			1	.70*
21							-.16*	.11							.12					.70*	1

Correlation Coefficients listed are significant at the .05 level unless denoted with an asterisk (*) for the .01 level. *Legend*: 1=ELA 2008 CST Score, 2=School Site Enrollment, 3=SES Participation, 4=Gender, 5=Ethnicity, 6=Disability, 7=Grade 3, 8=Grade 4, 9=Grade 5, 10=Grade 6, 11=Parent Not High School Grad, 12=Parent High School Grad, 13=Parent College Educated, 14=Decline Educ. Report, 15=Student English Speaker, 16=Student Spanish-English Speaker, 17=Student English Learner, 18=Student Reclassified Fluent Learner, 19=Student Free/Reduced Lunch Participant, 20=School Mobility, 21=District Mobility

Chapter 4

RESULTS AND DISCUSSION

In this chapter, I detail the statistical analysis I conducted to estimate the impact of Supplemental Educational Services (SES) on standardized test scores in English-language Arts (ELA) and mathematics using the data set previously described. I present three prospective functional forms for each dependent variable and consider what validity tests and adjustments are necessary to correct for potential multicollinearity and issues of heteroskedasticity. Selection of the appropriate regression equation will be determined based upon an analysis of the factors found to be statistically significant within a probability of 90 percent or greater, the positive or negative directions of the variable coefficients compared with the hypotheses, and a comparison of the three adjusted R^2 coefficients of determination revealed by each prospective functional form. Following the statistical analysis are the presentation and discussion of my findings from interviews with school administrators regarding their perceptions of the effect of the SES program on students' academic performance.

Selecting a Functional Form

If the independent or explanatory variables in the regression equation are specified correctly, the regression will yield a slope coefficient for each variable that provides an estimated quantitative effect of a one-unit change in a specific explanatory variable on the 2009 CST score, the dependent variable, as if all other explanatory variables were held constant. This regression estimation technique, known as Ordinary Least Squares, accounts for random errors in a data set by calculating the estimated slope coefficients to minimize the difference between the actual quantitative effect on each observation and the estimated effect produced from the aggregate data. Minimizing this difference (the residual) allows the researcher to consider whether the equation is predictive of similar outcomes among larger populations (Studenmund, 2006).

For this data set, the regression analysis includes a comparison of three prospective functional forms. Explanatory variables specified correctly may appear to be insignificant or to have an unexpected sign if the researcher selects an inappropriate functional form for the equation. The simplest functional form is a linear regression, wherein plotting the dependent variable as a function of the independent variables will generate a straight line (depicting a linear relationship). For example, in interpreting the simple linear regression for this study, the result provides an estimate of how the standardized test score of the *i*th student would increase or decrease according to the aggregate value of the equation coefficients multiplied by the value of their associated variables. Holding all variables constant, the equation will isolate the effect of SES participation and reveal statistically whether or not the effect is significantly different from zero.

An alternative functional form is the log-linear form in which the natural logarithm of the dependent variable is a function of the unlogged or linear independent variables. The use of this function describes an impact of the independent variable on the dependent variable in percentage terms such that if the value of explanatory variable X increases by one unit, the dependent variable will change in percentage terms that is measured by 100 multiplied by the coefficient for that X variable. A third alternative functional form is the log-log form using the logarithm of the dependent variable and all non-dichotomous explanatory variables so that the regression is nonlinear in the variables but still linear in the coefficients (Studenmund, 2006, p. 213). In the log-log equation, the coefficients for the independent variables indicate a percentage change in the dependent variable when the value of a specific independent variable increases by one percent.

Student proficiency is tested separately for English-language arts (ELA) and mathematics, and the 2009 scores for each test serve as the dependent variables for two separate

regression equations. I address the selection of the appropriate functional form for the 2009 CST for ELA first.

Dependent Variable of 2009 CST Scores for ELA

Table 7 includes the results of the three alternative functional forms for the dependent variable of 2009 CST scores for ELA. Following the comparison of these initial regression results, I discuss the tests and adjustments for multicollinerity and heteroskedasticity prior to presenting the results of the equation I selected for this dependent variable. I present the functional forms reviewed for the dependent variable of mathematics scores in the second portion of this chapter.

Table 7: Functional Forms for Regression of 2009 CST Scores for ELA

Variables	Lin-Lin	Log-Lin[!]	Log-Semi-Log[!]
Independent Variables: Student Academic Aptitude			
2008 CST Score for ELA	.737*** (.040)	.002*** (.000)	.724 (Ln)*** (.041)
Independent Variables: School Environment			
School Site A	2.906 (4.125)	.010 (.013)	.008 (.013)
SES Participant	-1.772 (4.307)	-.006 (.014)	-.008 (.014)
Male	-8.794*** (3.250)	-.025** (.0103)	-.027*** (.010)
Hispanic	-7.969 (6.118)	-.020 (.019)	-.020 (.019)
Disability	-36.486*** (13.095)	-.132*** (.041)	-.131*** (.041)
Grade 3	-29.730*** (4.729)	-.094*** (.015)	-.098*** (.015)
Grade 5	-22.982*** (4.609)	-.069*** (.015)	-.073*** (.015)
Grade 6	1.165 (4.853)	.008 (.015)	.003 (.015)

Table 7: Functional Forms for Regression of 2009 CST Scores for ELA – Continued

Variables	Lin-Lin	Log-Lin[!]	Log-Semi-Log[!]
Independent Variables: Social Factors			
Free or Reduced Lunch Program	-46.796* (30.705)	-.141* (.097)	-.162** (.097)
Parent Not High School Grad	-2.076 (5.768)	-.009 (.018)	-.008 (.018)
Parent High School Grad	-.493 (5.388)	-.004 (.017)	-.003 (.017)
Parent College Experience	9.970** (5.901)	.029* (.019)	.030* (.019)
English Second Language	.113 (8.124)	-.006 (.026)	.001 (.026)
Fluent-English Reclassified	11.635** (5.499)	.035** (.017)	.034** (.017)
English Learner	-3.696 (4.894)	-.013 (.015)	-.010 (.016)
School Mobility	2.615 (4.791)	.010 (.015)	2.6142 (4.9412)
District Mobility	1.826 (5.151)	.005 (.097)	.007 (.015)
R-Squared	.7142	.7047	.7043
Adj. R-Squared	.6975	.6873	.6869
Observations	325	325	325
Total Variables Significant	5***, 2**, 1*	4***, 2**, 2*	5***, 2**, 1*

Statistical significance ≥ 90% indicated by *, ≥ 95% by **, and ≥ 99% by ***.
Reference variables include Grade 4, No Education Reported by Parent, and English Only Student. Standard errors are in parentheses below the regression coefficients.
[!] The log forms of only the dependent/non-dichotomous variables are used.

The results included in Table 7 above relate to the dependent variable of ELA test scores as a function of explanatory variables classified as Student Aptitude, School Environment, and Social Factors. The coefficients and the standard errors for each explanatory variable measure the effect of the respective variables on the dependent variable and reveal whether the effect is significantly different from zero for this data set. Table 7 reveals there are explanatory variables affecting ELA test scores that are significant across all three functional forms. For instance, note that the variable for the 2008 CST score has a positive effect on the dependent variable of 2009 CST scores for ELA, with at 99 percent probability that the effect is significantly different from

zero. The variable identifying a student who is reclassified as English fluent estimates a positive effect across all forms, with 95 percent probability, compared to the variables denoting a student who speaks English only, one who is Spanish-English bilingual, or one who is learning but not yet English fluent. The variable defining a student whose parent has attended college has a positive effect with 95 percent probability according to the Linear form and 90 percent per the Log-Lin and Log-Semi-Log forms, when compared to a student whose parent graduated high school, did not graduate, or did not report education level.

There also are coefficients found to have a significant but negative effect on the 2009 ELA scores across all functional forms. For instance, the regression equations estimate that test scores will be lower if a student is disabled compared to one who is non-disabled, or if a student is in grade 3 or 5 compared to one in grade 4. The Linear and Log-Semi-Log forms estimate with 99 percent probability that a student who is male compared to female will have a lower 2009 CST score; the same is found with 95 percent probability per the Log-Lin form. A student who participates in the free or reduced lunch program is estimated to perform lower than a non-participating student does, with a 90 percent probability according to the Linear and the Log-Linear but with 95 percent probability per the Log-Semi-Log form.

The variables denoting school site enrollment and SES participation indicate a null effect on 2009 ELA scores across all functional forms. In other words, because the test statistic (p-value) indicates the coefficients for these variables are not significantly different from zero, I cannot reject the null hypothesis (that the coefficient would be equal to zero). A null effect on ELA scores is also estimated for the following variables: a student in grade 6 compared to one in grade 3, 4, or 5; one whose parent graduated high school, did not graduate high school, or did not report education compared to one whose parent has college experience; a student who is Spanish-English bilingual or one learning but not yet English fluent compared to one who speaks English

only or is reclassified as English fluent; a student who attended a different school prior to the 2008 CST; or one who attended a different district prior to the 2008 CST.

Although the initial findings estimate that the coefficient for SES Participation is not significantly different from zero and has a null effect on the dependent variable of 2009 CST scores for ELA, the next step in the analysis is to determine the most appropriate functional form for the regression equation. In examining all three forms, a comparison of the number of statistically significant variables and the goodness of fit for each equation pursuant to the R^2 coefficients of determination indicates that the linear regression is the most appropriate form to quantify the factors that affect 2009 CST scores for ELA. The linear form reveals there are five variables out of 21 found to be statistically significant within a probability of 99 percent, and there are two additional variables significant within a probability of 95 percent, for seven variables that are significant within a marginal significance level of .05 or less. The R^2 measure for the linear regression is highest at .7142, which implies that this equation explains 71.42 percent of the variations in the 2009 CST scores for ELA. In addition, comparing the directional effect of those variables that are statistically significant at the level of 95 percent probability or better shows consistency with the alternative hypotheses previously stated in Chapter 3, and presented below in Table 8.

Table 8: Alternative Hypothesis Compared with Linear Regression Results

Independent Variable	Null Hypothesis	Alternative Hypothesis	Direction of Coefficient
2008 CST Score ELA	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive
Male Gender	$H_0 : \beta \geq 0$	$H_A : \beta < 0$	Negative
Disability Status	$H_0 : \beta \geq 0$	$H_A : \beta < 0$	Negative
Grade 3	$H_0 : \beta = 0$	$H_A : \beta \neq 0$	Negative
Grade 5	$H_0 : \beta = 0$	$H_A : \beta \neq 0$	Negative
Parent College Experience	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive
Fluent-English Reclassified	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive

Having selected the appropriate functional form, further evaluation of the equation for multicollinearity and heteroskadacity is discussed below, which is then followed by presentation of the final regression results for the ELA dependent variable.

Testing for Multicollinearity in the 2009 CST for ELA Equation

The Ordinary Least Squares regression technique is based upon various logical assumptions, one of which is that no independent variable is a perfect linear function of any other independent variable (Studenmund, 2006, p. 245). Multicollinearity hinders the regression equation from distinguishing the statistical effect of independent variables upon the dependent variable due to correlation. Detecting multicollinearity takes into account that it is a common occurrence in many statistical analyses and that the degree of multicollinearity is the primary concern in addressing the issue and testing for it.

The calculated variance inflation factor (VIF) provides an index of how multicollinearity increases the variance of an estimated coefficient, and a VIF value of equal to or greater than 10 indicates that multicollinearity exists for a pair of variables. The VIF values reported below in Table 9 reveal no evidence of multicollinearity among the variables. This finding is acceptable given the number of statistically significant variables and the R^2 statistic that indicates the regression equation explains 71 percent of the variation in the dependent variable.

In discussing the issue of multicollinearity, it is useful to reiterate that, of the 21 independent variables in each equation, eight variables are dichotomous or dummy variables where only the reference variable is included in the equation, expressed with a value of one if affirmed. For example, the dummy variable represents a student who either attended School Site A, or is a SES Participant, male, Hispanic, disabled, previously attended a different school, or attended a different district, respectively. Variables such as grade level, parent education level, and language fluency are categorical variables and their associated subcategories are

dichotomous. When running the regression command with all explanatory variables listed in the equation, the STATA outcome reports the following variables as automatically omitted due to collinearity: grade 4 (which is 21 percent of the 325 observations), Parent Education Not Reported (11 percent of observations), and English Only Student (25 percent of observations). Thus, the estimated coefficients for these categorical variables (that is, the effect and magnitude of each specified grade level, parent education level, and language fluency) are evaluated based upon a comparison with the relevant associated variable that STATA reported as omitted due to collinearity.

Testing for Heteroskedasticity in the 2009 CST for ELA Equation

The Ordinary Least Squares regression technique assumes also that the variance of the error term is constant in a correctly specified equation (Studenmund, 2006, p. 346). When the error term variance is not constant for all observations, but changes from observation to observation, heteroskedasticity exists in the data set. Heteroskedasticity frequently exists with data sets comprised of widely ranging observed values of the dependent variable. Among the consequences of not correcting for heteroskedasticity is the potential for unreliable hypothesis testing due to biased standard error statistics (Studenmund, 354). Testing for heteroskedasticity involves various different methods such as the Park test and the White test described by A. H. Studenmund (2006, p. 358-362), as well as the STATA functional commands known as the Breusch-Pagan/Cook Weisberg test for heteroskedasticity and Szroeter's test for homoskedasticity.

To evaluate heteroskedasticity in the 2009 CST scores for ELA linear regression, I first consider a null hypothesis that the variance in the error term is constant. The Breusch-Pagan test for heteroskedasticity calculated with STATA reports the Prob > chi2 score is equal to .0036, which is within the rule that the chi2 score must be equal to or less than .10 in order to reject the

null hypothesis of constant variance. Referring to the descriptive statistics in Table 4 of Chapter 3, the potential existence of heteroskedasticity is consistent with the observed variation in the values of the dependent variable, which range from a minimum score of 222 to a maximum score of 476 for the ELA standardized test. Szroeter's test for homoskedasticity, with the null hypothesis that the variance is constant, also reveals the likelihood that heteroskadisticity exists, and indicates that the standard errors are over estimated for the following independent variables: SES Participation, Grade 3, Parent High School Graduate, and Parent College Graduate. Results of the linear regression equation using heteroskedasticity-corrected standard errors, which decreases the probability of significance, are shown in Table 9 below, with the VIF scores confirming no evidence of multicollinearity.

Final Regression Results for 2009 CST Scores for ELA

The final regression results presented in Table 9 include the variable coefficients and standard errors corrected for heteroskedasticity. The analysis quantifies the factors that impact student performance on the California standardized test for ELA.

Table 9: Final Regression Statistics for Effect on 2009 CST Scores for ELA (HC)

Variables	Linear (Uncorrected)	VIF Uncorrected	Best OLS Fit Linear (HC)	VIF Corrected
Independent Variables: Student Academic Aptitude				
2008 CST Score for ELA	.724*** (.041)	1.57	.724*** (.039)	1.57
Independent Variables: School Environment				
School Site A	2.906 (4.125)	1.33	2.906 (4.199)	1.33
SES Participant	-1.772 (4.307)	1.20	-1.772 (4.965)	1.20
Independent Variables: Social Factors				
Male	-8.794*** (3.250)	1.12	-8.794* (3.387)	1.12
Hispanic	-7.969 (6.118)	1.64	7.969 (6.796)	1.64
Disability	-36.486*** (13.095)	1.32	-36.486*** (7.205)	1.32
Grade 3	-29.730*** (4.729)	1.85	-29.730*** (4.911)	1.85
Grade 5	-22.982*** (4.610)	1.85	-22.982*** (4.533)	1.87
Grade 6	1.165 (4.853)	1.79	1.165 (4.645)	1.80
Free or Reduced Lunch Program	-46.796* (30.705)	1.23	-46.796 (40.593)	1.23
Parent Not High School Grad	-2.0756 (5.768)	2.55	-2.0756 (5.904)	2.55
Parent High School Grad	-.4932 (5.388)	2.91	-.4932 (5.331)	2.91
Parent College Experience	9.970** (5.901)	2.89	9.970* (6.325)	2.89
English Second Language	.113 (8.124)	1.39	.113 (.8.409)	1.39
Fluent-English Reclassified	11.635** (5.499)	1.93	11.635** (5.342)	1.93
English Learner	-3.696 (4.894)	2.54	-3.696 (5.260)	2.54
School Mobility	2.615 (4.791)	2.36	2.615 (4.816)	2.36
District Mobility	1.826 (5.151)	2.14	1.826 (5.384)	2.14

Table 9: Final Regression Statistics for Effect on 2009 CST Scores for ELA (HC)
- Continued

Variables	Linear (Uncorrected)	VIF Uncorrected	Best OLS Fit Linear [(HC)]	VIF Corrected
R-Squared	.7142	-	.7142	-
Observations	325	-	325	-
Total Variables Significant	8	-	7	-
Statistical significance ≥ 90% indicated by *, ≥ 95% by **, and ≥ 99% by ***. Reference variables include Grade 4, No Education Reported by Parent, and English Only Student. Robust standard errors are in parentheses below the regression coefficients. [HC] Heteroskedasticity Corrected Regression Coefficients				

Based on the 325 observations in the data sample and included in the linear regression, I cannot reject the null hypothesis that a student's participation in the SES program has no effect on that student's 2009 California Standardized Test score for English-Language Arts. The estimated coefficient of -1.772 for the SES participation variable has a standard error of 4.965 and a t-score of -0.36. The reported p-value for the SES participation variable is .681, which indicates there is 31 percent probability that this statistical result could be due to chance.

Three factors in the regression equation have a significant positive effect on 2009 test scores for ELA in this sample. The factors include a student's score on the 2008 CST for ELA; the variable identifying a student whose parent has education at the college level compared to one whose parent did not report education level; and the variable identifying a student who was reclassified as English fluent prior to 2008 compared with a student who speaks English only.

Several factors in the equation have a significant negative impact on the test scores of students in this sample, including gender, disability status, and grades 3 or 5. For example, the variable denoting gender estimates that the 2009 test score of a male student will be 8.794 lower than a female student's score. A student with a disability is estimated to score 36.386 lower than one who is not disabled; a student in grade 3 is estimated to score 29.730 points lower; and one in grade 5 is estimated to score 22.982 lower when each are compared with a student in grade 4.

The regression analysis indicates there is a null effect on 2009 test scores related to SES participation. The regression also indicates a null effect on the 2009 test scores related to the variable that identifies a student who attended School Site A versus Site B. Additional variables that have a null effect according to the regression equation include those that identify a student who: is Hispanic compared with non-Hispanic; is in grade 6 compared to grade 4; has a parent who did or did not graduate high school compared to one whose parent did not report education level; is Spanish-English bilingual compared to English fluent only; participated in the free-or-reduced lunch program; attended another school prior to 2008; or attended another district prior to the 2008.

A comparison of the mean of the observed 2009 CST scores for ELA with the mean of the predicted 2009 CST scores based on the regression equation indicates no significant difference in the two statistics. On average, students' scores appear to be consistent with the estimated score predicted by the regression coefficients for this sample. The standard deviation of the observed scores, however, is slightly greater than that for the predicted scores, which indicates that after fitting the estimated coefficients to the equation, the distribution of the estimated scores is somewhat closer to the predicted mean.

Table 10: Comparison of Mean CST Scores for English-Language Arts

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Observed 2009 CST for ELA	325	323.7538462	50.32641	222	476
Predicted 2009 CST for ELA (Yhat)	325	323.7538465	42.53158	223.8793	495.4529

After fitting the estimated coefficients to the equation to obtain the predicted 2009 CST scores for ELA, I calculated the difference between each observed score and each predicted score,

on a case-by-case basis, and found that the mean of this difference is -3.29, with a standard deviation of 26.904. This indicates that the student scores observed in this sample are, on average, 3.29 points lower than the score predicted by the linear regression. The differences in observed versus predicted scores ranges from 74.5 points lower than the predicted score to 97.7 points more than the predicted score.

Dependent Variable of 2009 CST Scores for Mathematics

Table 11 below includes the results of the three alternative functional forms for the dependent variable of 2009 CST scores for mathematics. The additional tests and adjustments for multicollinearity and errors related to heteroskedasticity precede the results related to this dependent variable.

Table 11: Functional Forms for Regression of 2009 CST Scores for Mathemathics

Variables	Linear	Log-Lin[!]	Log-Semi-Log[!]
Independent Variables: Student Academic Aptitude			
2008 CST Score for Math	.637*** (.040)	.002*** (.000)	.634*** (.040)
Independent Variables: School Environment			
School Site A	9.548 (6.764)	.030 (.019)	.031 (.019)
SES Participant	-8.165 (7.098)	-.024 (.020)	-.026 (.020)
Independent Variables: Social Factors			
Male	-4.907 (5.366)	-.014 (.015)	-.013 (.015)
Hispanic	-1.963 (10.123)	-.005 (.029)	-.010 (.029)
Disability	-17.157 (21.583)	-.059 (.062)	-.054 (.062)
Grade 3	15.272** (7.723)	.036 (.022)	.040* (.022)
Grade 5	-19.103** (7.482)	-.064*** (.021)	-.065*** (.021)
Grade 6	23.252*** (8.224)	.069*** (.024)	.074*** (.024)
Free or Reduced Lunch Program	-11.565 (50.525)	-.068 (.145)	-.049 (.145)
Parent Not High School Grad	11.233 (9.477)	.036* (.027)	.041* (.027)
Parent High School Grad	14.010* (8.877)	.041* (.025)	.045** (.025)
Parent College Experience	17.534** (9.720)	.050** (.028)	.056** (.028)
English Second Language	-3.456 (13.333)	-.0128 (.038)	-.004 (.038)
Fluent-English Reclassified	13.400* (9.152)	.039** (.026)	.042* (.026)
English Learner	-12.563* (7.844)	-.041** (.023)	-.038** (.022)
School Mobility	-4.563 (7.891)	-.010 (.023)	-.015 (.023)
District Mobility	-.377 (50.525)	-.009 (.024)	-.006 (.024)

Table 11: Functional Forms for Regression of 2009 CST Scores for Math - Continued

Variables	Linear	Log-Lin[!]	Log-Semi-Log[!]
R-Squared	.5895	.5982	.5993
Adj. R-Squared	.5653	.5745	.5757
Observations	325	325	325
Total Variables Significant	2***, 3**, 3*	3***, 3**, 2*	3***, 3**, 3*
Statistical significance ≥ 90% indicated by *, ≥ 95% by **, and ≥ 99% by ***. Reference variables include Grade 4, No Education Reported by Parent, and English Only Student. Standard Errors are in parentheses below the regression coefficients. [!] The log forms of only the dependent/non-dichotomous variables are used.			

The results included in Table 10 show that there are explanatory variables with statistically significant coefficients affecting math test scores across all three functional forms based upon 90 percent or better probability that the effect is greater than zero. For instance, the coefficient for the 2008 CST score is positive and statistically significant with 99 percent probability, as is the variable identifying a student enrolled in grade 6, which estimates higher performance on the standardized test for math compared to a student in grade 3, 4, or 5. A student whose parent has college experience is estimated with 95 percent probability to perform better than one whose parent does not have college or did not report education level. According to the Linear and Log-Lin forms, a student whose parent graduated high school is estimated at 90 percent probability to perform better than one whose parent has not graduated or did not report education, which increases to a probability of 95 percent in the Log-Semi-Log form. A student in grade 3 compared to grades 4, 5, or 6, is estimated to perform better at 95 percent probability according to the Linear, which decreases to 90 percent in the Log-Semi-Log form and is found to be insignificant per the Log-Lin form.

Additional variables significant across all functional forms that are negative include those denoting a student enrolled in grade 5 compared to grade 3, 4, or 6, and a student who is learning but not yet English fluent compared with one who speaks English only, is Spanish-English bilingual, or reclassified as English fluent.

Independent variables categorized to quantify school environment (School Site and SES Participation) are found to have a null effect on the dependent variable of 2009 CST scores for math across all functional forms. The results are similar for several of the variables quantifying social factors. For instance, across all functional forms there is no impact on the test scores that is different from zero for a student who is male compared to female, Hispanic or non-Hispanic, disabled or non-disabled. Participation or non-participation in the free-or-reduced lunch program is not shown to have an effect on the math tests that is significantly different from zero, which is shown also for a student who is Spanish-English bilingual compared with a student who speaks only English, one who is reclassified as English fluent, or one who is learning English. The mobility variables that identify a student who attended another school or district prior to 2008 also have a null effect on the dependent variable across all functional forms.

The same criterion used to select the ELA regression equation is applied to determine the appropriate functional form to evaluate the 2009 CST scores for math. Comparing the number of statistically significant variables, the Log-Lin and Log-Semi-Log functional forms identify three variables that are significant at 99 percent probability. At the level of 95 percent probability, each of these two forms result in three additional significant variables, and at the 90 percent probability, the Log-Log form shows there are and additional three significant variables. The R^2 coefficient indicating the goodness of fit is greater under the Log-Semi-Log form with 59.93 percent of the variation in the dependent variable explained by this regression compared to 58.95 percent explained by the Linear or 59.82 percent by the Log-Lin forms. I, therefore, select the Log-Log form as the best fit for the regression equation of the 2009 CST score form mathematics. Comparing the directional effect for the variables that are significantly different from zero at the 90 percent probability or better shows consistency with the alternative hypotheses previously stated in Chapter 3. Table 12 below includes the hypothesis and results.

Table 12: Alternative Hypothesis Compared with Log-Semi-Log Regression Results

Independent Variable	Null Hypothesis	Alternative Hypothesis	Direction of Coefficient
2008 CST Score Math	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive
Grade 3	$H_0 : \beta = 0$	$H_A : \beta \neq 0$	Positive
Grade 5	$H_0 : \beta = 0$	$H_A : \beta \neq 0$	Negative
Grade 6	$H_0 : \beta = 0$	$H_A : \beta \neq 0$	Positive
Parent Not High School Graduate	$H_0 : \beta \geq 0$	$H_A : \beta < 0$	Positive!
Parent High School Graduate	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive
Parent College Experience	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive
Fluent-English Reclassified	$H_0 : \beta \leq 0$	$H_A : \beta > 0$	Positive
English Learner	$H_0 : \beta \geq 0$	$H_A : \beta < 0$	Negative
! Directional effect of coefficient is not consistent with alternative hypothesis.			

Testing for Multicollinearity in the 2009 CST Scores for Math Equation

The same purpose applies with regard to testing for multicollinearity of the independent variables as described and carried out for the ELA scores. To reiterate, if multicollinearity exists, the regression equation will not discern the statistical effect of the independent variables upon the dependent variable due to correlation. The calculated variance inflation factor (VIF) is an index of how much the standard error is increased due to correlation with other independent variables, with a value of equal to or greater than 10 indicative of multicollinearity. The VIF values reported below in Table 12 reveal no evidence of multicollinearity among the variables. This finding is acceptable given the number of statistically significant variables and the R^2 statistic that indicates the regression equation explains 59.93 percent of the variation in the dependent variable.

In discussing the issue of multicollinearity, it is useful to review the fact that, of the 21 independent variables in the equation for 2009 CST scores for math, eight variables are dichotomous, or dummy, expressed with a value of one if affirmed. For example, a respective value of 1 is assigned to a student who attended School Site A, is a SES Participant, male, Hispanic, disabled, or previously attended a different school or attended a different district.

Variables such as grade level, parent education level, and language fluency are categorical variables and their associated subcategories are dichotomous. When running the regression command with all explanatory variables listed in the equation, the STATA outcome reports the following variables as automatically omitted due to collinearity: grade 4 (which is 21 percent of the 325 observations), Parent Education Not Reported (11 percent of observations), and English Only Student (25 percent of observations). Thus, I evaluate the estimated coefficients for these categorical variables (that is, the effect and magnitude of each specified grade level, parent education level, and language fluency) based upon a comparison with the relevant associated variable, which STATA reported as omitted due to collinearity.

Testing for Heteroskedasticity in the 2009 CST Scores for Math Equation

As discussed earlier with regard to the ELA equation, when the variance of the error term is not constant for all observations, but changes from observation to observation, heteroskedasticity exists in the data set. Heteroskedasticity frequently exists when there is a wide range of observed values of the dependent variable, which can result in unreliable hypothesis testing due to biased standard error statistics (Studenmund, 354). For the 2009 CST for math dependent variable, I again applied the Breusch-Pagan/Cook Weisberg test for heteroskedasticity and Szroeter's test for homoskedasticity in the equation for the 2009 math scores dependent variable.

The Breusch-Pagan test for heteroskedasticity calculated with STATA results indicate evidence of heteroskadisticity, which is consistent with the wide range of observed values in the dependent variable with scores ranging from a minimum of 182 to a maximum of 600 (refer to the descriptive statistics reported in Table 4, Chapter 3). Szroeter's test for homoskedasticity, with the null hypothesis that the variance is constant, also reveals the likelihood that heteroskadisticity exists, and indicates that the standard errors are over estimated for the

following independent variables: the logarithmic transformation of the 2008 CST Scores for Math; Grade 3; and Parent with College Experience. In the next section, I present and discuss the final results of the linear regression equation for the 2009 CST scores for mathematics.

Final Regression Results for the 2009 CST Scores for Mathematics

See Table 13 below for the variable coefficients and robust standard errors estimated with a linear regression equation using the Log-Semi-Log functional form. The dependent variable is the logarithmic transformation of the 2009 CST scores for mathematics; the independent variables are the logarithmic transformation of the 2008 CST scores for mathematics, along with the remaining dichotomous variables as listed. The coefficients for the independent variables indicate a percentage change in the dependent variable when the value of the subject independent variable increases by one percent, holding all other variables constant.

Based on the regression results, and similar to the results regarding the ELA test scores, I cannot reject the null hypothesis that the key independent variable, participation in the SES program, has no statistical effect on 2009 California Standardized Test score for mathematics. The estimated coefficient of -.029 for the SES participation variable has a standard error of .023 and a t-score of -1.27. The reported p-value for the SES participation variable is .205, which indicates there is 80 percent probability that this statistical result could be due to chance.

Several factors in the equation have significant positive impact on test scores. These include the variables identifying a student in grade 6 versus grades 4, one whose parent graduated high school or has college experience compared to one whose parent did not report educational level, or one who was reclassified as English fluent prior to the 2008 CST test in mathematics compared to one who is English fluent only.

Table 13: Final Regression Statistics for Effect on 2009 CST Scores for Mathematics (HC)

Variables	Log-Semi-Log[!] Uncorrected	VIF for Uncorrected	Log-Semi-Log[!] Corrected	VIF for Corrected
Independent Variables: Student Academic Aptitude				
2008 CST Score for Math	.634*** (.040)	1.38	.634*** (.046)	1.38
Independent Variables: School Environment				
School Site A	.031 (.019)	1.32	.031 (.021)	1.32
SES Participant	-.029 (.020)	1.20	-.029 (.023)	1.20
Independent Variables: Social Factors				
Male	-.013 (.015)	1.12	-.013 (.016)	1.12
Hispanic	-.010 (.029)	1.65	-.010 (.030)	1.65
Disability	-.054 (.062)	1.32	-.054 (.048)	1.32
Grade 3	.040* (.022)	1.82	.040 (.025)	1.82
Grade 5	-.065*** (.021)	1.80	-.065*** (.020)	1.80
Grade 6	.074*** (.024)	1.91	.074*** (.021)	1.91
Free or Reduced Lunch Program	-.049 (.145)	1.23	-.049 (.111)	1.23
Parent Not High School Grad	.041* (.027)	2.54	.041* (.028)	2.54
Parent High School Grad	.045** (.025)	2.91	.045** (.026)	2.91
Parent College Experience	.056** (.028)	2.89	.056** (.030)	2.84
English Second Language	-.004 (.038)	1.37	-.004 (.041)	1.37
Fluent-English Reclassified	.042* (.026)	1.97	.042* (.027)	1.97
English Learner	-.038** (.022)	2.41	-.038* (.024)	2.41
School Mobility	-.015 (.023)	2.36	-.015 (.022)	2.36
District Mobility	-.006 (.024)	2.14	-.006 (.023)	2.14

Table 13: Final Regression Statistics for Effect on 2009 CST Scores for Mathematics (HC) - Continued

Variables	Log-Semi-Log[!] Uncorrected	VIF for Uncorrected	Log-Semi-Log[!] Corrected	VIF for Corrected
Independent Variables: Social Factors				
R-Squared	.5993	-	.5993	-
Adj. R-Squared	.5757	-	-	-
Observations	325	-	325	-
Total Variables Significant	8	-	8	-

Statistical significance ≥ 90% indicated by *, ≥ 95% by **, and ≥ 99% by ***.
Reference variables include Grade 4, No Education Reported by Parent, and English Only Student. Robust standard errors are in parentheses below the regression coefficients.
[!] The log forms of only the dependent/non-dichotomous variables are used.

The factors predicted to have a significant negative effect on the dependent variable in this sample include the variables identifying a student who is in grade 5 compared to a student in grade 4, and one who is learning English but not yet classified as fluent compared to a student who is fluent in English or reclassified as fluent.

The final regression equation, corrected for heteroskadasticity, estimates a null effect on the 2009 CST scores for the variable related to SES participation as indicated earlier. Several additional independent variables in the regression that have a null effect on the 2009 CST score for mathematics as well. These include the variables that identify a student who is: enrolled at School Site A compared with School Site B; male compared to female; Hispanic compared to non-Hispanic; disabled; in grade 3 compared to one enrolled in grade 4, 5, or 6; Spanish-English bilingual compared with a student who speaks English only; participating in the free-or-reduced lunch program; a student who attended another school prior to 2008; one who attended another district prior to the 2008.

A comparison of the mean of the observed 2009 CST scores for math with the mean of the predicted 2009 CST score based on the regression equation indicates no significant difference in the two statistics. On average, students' scores appear to be consistent with the estimated score

predicted by the regression coefficients for this sample. The standard deviation of the observed scores, however, is slightly greater than that for the predicted scores, which indicates that after fitting the estimated coefficients to the equation, the distribution of the estimated scores is somewhat closer to the predicted mean.

Table 14: Comparison of Mean CST Scores for Mathematics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Observed 2009 CST for Math	325	5.791611	.2005062	5.204007	6.39693
Predicted 2009 CST for Math (Yhat)	325	5.791611114	.1552146	5.335526	6.229956

After fitting the estimated coefficients to the equation to obtain the predicted 2009 CST scores for math, I calculated the difference between each observed score and each predicted score, on a case-by-case basis, and found that the mean of this difference is .0008121, with a standard deviation of .1277712. This indicates that the student scores observed in this sample are, on average, .0008121 percent greater than the score predicted by the linear regression. In addition, the differences in observed versus predicted scores ranges from a score that is .360908 percent lower than the predicted score, to a score that is .3597789 percent greater than the predicted score.

Findings from Interviews with School District Administrators

In April 2012, I conducted three one-on-one interviews with school district administrators who were familiar with the Program Improvement schools of this study and the SES program the district is required to offer to students under NCLB. I asked each individual to share his or her perceptions about the mandated tutoring. The questions I asked focused on perceptions about whether the SES program contributes to an improvement in academic performance or has an

effect on student test scores, whether students and parents who participate in the program appear satisfied with the tutoring services, and other questions related to challenges in delivering the SES program and compliance with the mandate. Refer to Appendix B for the interview protocol.

Responses Regarding Perceptions of Effectiveness

The interview responses from administrators revealed that school officials are uncertain or doubtful as to whether the SES program is effective as an intervention to improve academic proficiency due to lack of evidence. Each administrator discussed that there are probable barriers to meaningful evaluation of the SES intervention which are related to the significant variability in the tutoring curriculum, the different providers' teaching credentials and experience, the methods of delivering the tutoring services, the providers' assessment tools, the rate of student enrollment and participation, and the rate of individual student's attendance. Other than the annual standardized testing that occurs at the student's school site, no other system is in place for school officials to isolate, monitor, or analyze the effect of the tutoring program on the academic performance of those students who participate in the program. Also, because there is no coordination between the SES tutoring providers and school district administrators with regard to alignment of the tutoring curriculum with classroom or homework lessons, the administrators expressed doubt that the tutoring has a significant effect on the students' standardized test scores.

For example, administrators indicate, that a student who participates in SES may demonstrate some academic improvement according to the pre- and post-program assessments conducted and reported by the student's SES provider as required by the state. However, regardless of the SES provider's report, the administrators have no empirical evidence that the reported academic growth is a result of the tutoring, that tutoring improves a student's level of academically proficient, or even that the tutoring prompts significant improvement on standardized tests that are conducted at the students' school site.

In addition, to some extent, each administrator referred to the practical difficulty involved with measuring the effect of the SES tutoring program. One administrator discussed the objective of determining whether changes in a student's academic performance are attributable more or less to the SES tutoring, compared to the academic interventions implemented daily in the classroom environment (such as the Response to Intervention Model and the Student Study Team process, for example). Two of the school administrators reported that the academic interventions implemented in the classrooms are designed to meet the specific academic needs of the individual student within the scope of the standardized curriculum. Because school officials are not informed about the curriculum offered by the tutoring firms who provide the SES, they do not know whether it is aligned with the district's standardized curriculum. In addition, although the SES providers are required to meet with a student and parents to develop a learning plan with academic objectives, school officials or teachers are not involved in the development of the plan in any way other than to provide parents with a report of their student's standardized test scores and percentile proficiencies. The district does not monitor the SES providers' tutoring plans for participating students.

The administrators were of the belief that each of the SES providers offers different and distinctive curriculum, which confounds the capacity to compare tutoring programs or to evaluate the effect of the tutoring intervention. Each administrator reported that just as the SES firms do not interact with teachers, administrators, or other school district officials to coordinate or align tutoring curriculum with the state standardized curriculum, the district or the schools likewise do not provide information to any of the SES providers either about curriculum or a student's academic needs. The district requests, however, that SES tutoring services focus on English-language arts, mathematics, or both, based upon prior aggregate CST scores and the schools' progress toward meeting annual proficiency levels in those categories.

On the other hand, as indicated earlier, the district does provide parents with information about their child's standardized test scores, and SES providers are required by the district to meet with the parents and enrolled students to develop an instructional plan with learning goals and objectives. The SES providers are also required to report the results of instructional plans to the district upon completion of the tutoring program, which is when the pre- and post-program assessments are provided to the district (in exchange for payment from the district for the tutoring services rendered). The district does not analyze the SES assessment data provided by providers to evaluate the impact of tutoring on individual or aggregate test scores, and administrators believe that there are barriers and limitations to such analysis, which I have described above.

Perceptions Regarding Service Delivery and Compliance

The administrators each described the district and school processes in place to comply with the mandatory SES intervention and explained that the district office assumes the lead role in ensuring the delivery of the tutoring service to eligible students. The district actively promotes and encourages eligible students to participate in the SES program at the beginning of each academic year. Students who participate in the SES program receive tutoring after school, and although tutoring sessions are typically conducted off campus and participating students are responsible for their own transportation to tutoring sessions, some tutoring firms have facilities use permits with the district allowing them to deliver tutoring on the campus of the program improvement school after regular school hours. Officials or teachers at the student's school are likely to have no knowledge as to whether or not a student has chosen to participate in the SES program.

Regarding the district's process for complying with the SES mandate, for example, the program is offered to students by the district but it is provided by independent tutors or tutoring firms who are in compliance with the district's documentation requirements and who are

approved by the California Department of Education (CDE). The district assigns each student to their SES provider of choice and remits payment for the tutoring directly to the SES provider once a student exhausts the allotted maximum per student benefit established by the state, or at the end of the academic school year, whichever comes first, and upon receipt of SES provider's student attendance record and assessment report. The literature provided to parents by the district to announce the SES program advises parents to consider the SES provider's cost per hour charge for tutoring when making the request for a provider. Although the district pays the SES provider directly for the tutoring services, the variation in the hourly rate charged by each tutoring firm impacts the number of tutoring sessions a student may receive under the SES program. The administrators are of the opinion that the effect of tutoring may be increased or may be more beneficial to a student as the number of tutoring sessions increases.

When asked whether school district officials encourage student participation in the SES program, the administrators each responded affirmatively and explained that the schools host annual on-site tutoring fairs sponsored by the district, where the SES providers market their tutoring programs to parents and students. Following the annual tutoring fairs, parents may then opt to enroll their child in an approved tutoring program by submitting an application for the services to the district, which is typically collected by the student's school office. The application includes basic demographic information about the student and his or her family such as name, address, grade, and parent contact information, in addition to the parent's preference list of particular SES providers requested for the student's tutoring. The district processes the application and notifies parents directly of the SES provider assignment. Officials and teachers at the student's school are not in any other way involved in a parent's application for or selection of a tutoring program and may not have knowledge as to whether or not a student chooses to participate in the SES program.

In addition, the administrators advised me that when parents inquire about helping their children improve, teachers informally encourage the parents to contact the district office to find out if a student meets the criteria for SES participation. Eligibility is met if a student attends a Title I school and is from a low-income family as determined by participation in the free-or-reduced lunch program. Priority for placement in the SES program is given to low-achieving students according to performance on the standardized tests for ELA and math.

The general conclusions, therefore, regarding the district's compliance with the SES mandate and the delivery of the program are that there may be a need to improve communication between the district and the SES providers in terms of alignment of curriculum or monitoring of student progress. Other than district-sponsored tutoring fairs held on campus at the beginning of each academic year and the submittal by providers of their attendance records and assessment reports at the end of the year, there are no other communications between the district and the tutoring agencies or between the schools and the tutoring agencies. Although a parent may contact the district on occasion to request a change in SES assignment, according to the interview responses, the district does not monitor student attendance in the tutoring program for any reason other than to remit payment to the tutoring providers. Administrators or teachers at the student's school also do not monitor SES enrollment or attendance in any way and may only become aware of a student's participation in SES based upon informal communications with a student or parents.

Perceptions of Customer Satisfaction

The school district administrators have no empirical knowledge as to whether parents or students are satisfied with the SES program or the tutoring providers. The district does not conduct any type of survey to quantify participants' opinions or their assessment regarding their experience with the SES program. The same is true in terms of evaluating whether school district

officials and teachers are satisfied with the SES program. The administrators base their assessment of customer satisfaction on informal communications with parents, students, and teachers, as described below.

Each administrator reported occasionally receiving informal feedback from parents independently, which indicates a general opinion that the program is a valuable augmentation to a participating student's academic opportunities. Similarly, administrators report that a participating student may informally express an opinion to school officials or teachers related to their level of satisfaction with the program. Statements from students may vary widely, however, and range from a typical student reaction of dismay about having to attend tutoring sessions and spend afterschool time studying to a frequent student response that is more favorable about the tutoring but usually related to approval of the attendance incentives a student may receive from the tutoring provider, such as afterschool snacks.

I asked each administrator about whether, in his or her knowledge or opinion, teachers are satisfied with the SES program or the tutoring providers. One administrator responded that teachers are hesitant to have a positive judgment about the tutoring intervention program due to uncertainty about alignment of the tutoring curriculum and due to the lack of empirical evidence that the tutoring program improves proficiency or effects test scores. Another administrator responded that teachers are likely to be dissatisfied with the SES program related to their opinions that there are inadequate credential requirements for tutoring firms, that there is inadequate disclosure of individual tutor's credential qualifications, and that there is inadequate information regarding the quality of the tutoring curriculum implemented by the SES providers.

As for whether school district officials are satisfied with delivery or results of the SES program, one administrator responded in terms of the amount of Title I funds expended for the SES program, affirming that the district fully disburses the funds set aside in compliance with the

mandate. The administrator expressed uncertainty or doubt, however, that the per-student allocation fixed by the CDE is sufficient to provide enough tutoring to significantly affect the academic proficiency of participating students. In addition, the administrators were of the opinion that, although the district is able to offer the SES tutoring to all students who are interested and eligible to participate, the level of participation is relatively low. Those students who participate in the SES program may benefit from the tutoring in terms of improved proficiencies in ELA, math, or both. However, because the rate of participation in the SES program is so low, the school may continue in program improvement because the aggregate percentage rate of students who are proficient may continue to be too low to meet the annual yearly progress prescribed by the NCLB mandate. Two administrators expressed their opinions that, although the current level of funding for the SES program may be insufficient and unsatisfactory, additional funding should not be allocated until and unless the SES program is empirically proven to be an effective intervention to increase aggregate academic proficiency.

The interviews and discussions with the administrators provided qualitative details about implementation of the SES program at the local level. Some of the opinions and beliefs shared by the administrators are consistent with the statistical analysis of the data. For instance, the administrators' comments regarding the degree of difficulty in quantifying the isolated effect of SES participation from other school, social, or socio-economic factors that may affect student academic achievement correlates with the statistical findings that SES participation has an estimated null effect on the test scores observed in this sample. In addition, the interview comments regarding the rate of student participation in the SES program and the relative effect on aggregate proficiency levels are consistent with the data in this sample. Less than 20 percent of students in the sample observed for this study were SES participants, and there was no significant

change from one year to the next in the mean test scores to indicate aggregate academic improvement.

In the next chapter, I conclude this report with a proposed explanation about why the key explanatory variable of SES participation is insignificant in this study, and I offer suggestions for replicating the research in the future. I discuss the limitations of the data and the regression equation and offer a recommendation for expanding and applying the analysis as an evaluative tool to assess the performance of SES providers.

Chapter 5

CONCLUSION

The purpose of this study has been to measure the effect of the tutoring intervention provided to Title I students through the Supplemental Educational Services (SES) component of *No Child Left Behind Act* on academic achievement. I obtained and analyzed data from a typical suburban school district in Northern California, which included 325 observations of elementary school student test scores and demographic information. The results of the linear regressions I conducted indicate that the key explanatory variable of SES participation did not have a statistically significant effect on students' academic performance as measured by the 2009 California Standardized Tests in English-language Arts or mathematics.

In this chapter, I propose possible explanations as to why the SES participation variable is not statistically significant compared with other variables included in the study. I offer suggestions to expand the data and the equations in future studies to better account for the rate of attendance, program completion, particular SES providers, types of curriculum, and tutoring method. I discuss some implication of the regression analysis and consider how the findings may be applied to assess whether the SES provision is an appropriate use of Title I funds to promote academic proficiency among low-income students. I also describe the limitations of the regression equation and the data available to quantify the effect of SES participation on student achievement. Finally, I conclude this thesis with a recommendation for expanding this regression analysis systematically to evaluate the performance of SES tutoring firms as effective partners in the effort to provide access to quality education for all children.

Explanations and Suggestions Regarding Statistical Significance

In my final data set, with 325 observations of SES eligible students, only 61 of these students received at least one hour of tutoring through the SES program in the academic year

preceding the 2009 CST. Among these 61 students, the mean total time spent in tutoring during that period was 15.7 hours, with a median of 17.6 hours. Among the 325 total observations, the mean time spent in tutoring was only 3.46 hours, with a median of zero. These descriptive statistics reveal a potential bias in the distribution of the sample population in this study that would minimize an effect of the SES tutoring on aggregate student achievement.

For future research, if the data set were larger and included more observations of SES eligible students and a greater proportion of SES participants, I suggest that an interaction variable be included in the regression equation to better account for the percentage of the SES program completed by a student. The additional variable may improve the strength of the regression to estimate the SES impact and measure any cumulative effect of participation and attendance on test scores that is statistically significant.

As a key dichotomous variable in this study, the regression equation accounted only for whether or not an eligible student participated in the SES program, without further explanation regarding provider, tutoring curriculum, or delivery methods. District records used for this study indicate that students who participated in SES enrolled with one of thirteen SES providers. For the future, if replicating this study with a larger sample and greater proportion of SES participating students, I would suggest including one or more interaction variables in the regression analysis to account for the quantity of tutoring time as well as the type of SES provider, the nature of the curriculum (English-language arts or mathematics), and the method of tutoring sessions (such as small group, one-on-one, or internet based). The additional interaction variables may reduce potential bias and may yield a finding that the SES intervention has a significant effect among some types of SES providers, curriculum, or delivery method compared to other alternatives.

Implications of the Regression Analysis for Assessment of the SES Intervention

Expanding the analysis to include more observations or to include some or all of the variables described earlier may improve the statistical evaluation of the SES impact on student achievement at the local level. Disaggregating the data, on the other hand, also provides valuable information that exposes the technical challenges of measuring academic achievement and imposing sanctions according to the percentage of proficient students.

Based on school-wide data available from the CDE and the district, the two schools in this study have been in program improvement since 2004-05 in one case and since 2006-07 in the other; both having also failed to reach the state's API target for achievement growth twice between 2008-09 and 2010-11. The negative correlation coefficients found in this study for the variables designated as 2008 tests scores, ethnicity, and English fluency indicate an opposing relationship between each of these factors (see Tables 5 and 6 in Chapter 3). The accountability reports available through the California Department of Education website confirm that the intermittent success or failure of the schools to achieve AYP and to meet API is related to the low rate of proficiency among a majority of the student subgroups served by each school (students of Hispanic ethnicity, students who are English learners not yet fluent, and students who are socio-economically disadvantaged). Fundamental to the public education system in the United States is the idea that all children are entitled to have access and to receive an education. Yet the failure of student subgroups to demonstrate increasing levels of aggregate proficiency as prescribed under NCLB and according to standardized tests should not necessarily implicate the school as failing to provide quality educational services to these children.

The linear regression in this study indicates there is evidence of the negative impact of some social factors that are outside of the student's school environment, such as a student's disability status, a student's age, or whether the student's family is low income, on test scores.

And there is also evidence of the positive impact of other social factors outside of the student's school environment, such as a parent's reported education at the college level, on student test scores. Another important finding from the regression analysis is the estimated positive and significant effect on test scores related to the variable describing a student who was reclassified in 2008 from English Learner to fluent in English compared with a student who is fluent in English only. The positive effect of this factor may indicate that these students are benefiting academically from strategies to improve language fluency that are taking place in learning environments inside and outside of the classroom.

Labeling schools for program improvement interventions motivates schools and districts to implement intensive evidence-based strategies that expand and enhance the educational services they provide to low-achieving and at-risk students. However, program improvement interventions such as SES under NCLB, which requires districts to set aside 15 percent of Title I funds for students to attend tutoring services provided by private firms outside of the school environment, may effectively decrease the total funding resources of Title I schools. A reduction in total funding, at a time when student populations appear to require intensive educational services, may disadvantage the school's own program improvement efforts that perhaps serve more children than SES and be more effective if fully funded. It may be that the implementation of the SES component of NCLB lacks sufficient intensity or may be too fragmented to have a significant effect for improving academic performance of low-achieving students, particularly when compared to how the resources might be consolidated and used to implement research-based interventions in the schools and classrooms. Adding more rigorous assessment measures to the SES program would likely improve the Pareto efficiency of the program and may increase the impact of the intervention on participating students' academic performance.

Limitations of the Research Equation and Data

The finding that SES participation does not have a significant effect on the test scores of students in this sample is consistent with some of the findings in the previous studies conducted at school districts across the nation and published since the implementation of the intervention. The low rate of participation in the SES program among eligible students and the low attendance rate of students who did voluntarily enroll are details surrounding the data that may impose limitations on the conclusiveness of the regression results. The explanatory variables included in the equation pose limitations on the power of the regression to measure the SES effect to the extent that variables were restricted to student test scores and student demographic information received from the school district. As discussed above, expanding the size of the sample and the number of variables in the equation may reduce potential bias and strengthen the regression to measure the effect of SES participation, attendance, and program type on student achievement on test scores.

This study's data set consisted of only 325 observations, which is sufficient but relatively small for a regression equation with 21 independent variables. Conducting the regression analysis with a larger data set may reduce potential errors. Since the time I initially compiled the data set for this study in 2009 and 2010, two additional schools have entered program improvement status and are required to offer SES services as of 2011-12 or 2012-13. The additional population of students eligible for the SES program creates the opportunity to improve the regression analysis with a larger sample. The increased eligibility for the SES program may also produce an imperative to replicate this study to more conclusively measure the effect of SES as discussed earlier in this chapter.

Recommendations for Evaluating SES Providers

Examining the student-level data and the effect of the additional variables on students' test scores would certainly provide the states, school districts, and parents with useful assessment information about the quality of the SES tutoring firms and their capacity to raise students' academic achievement on standardized tests. The practice of allocating public Title I funds to private tutoring firms, without implementing more rigorous techniques to evaluate the effect of the outside tutoring compared to the classroom intervention programs creates inefficiencies at many levels and places program improvement schools and districts at greater disadvantage economically. The potential negative externalities related to Pareto efficiencies of the intervention ultimately affect all students who attend the schools within the district.

California regulations require tutoring vendors to demonstrate a record of increasing student's academic proficiency in order to maintain the firm's status as a vendor authorized to receive payment for the tutoring from the districts they serve. The various student-level learning plans and pre- and post-treatment assessments that the SES providers currently are required to submit to the districts should be standardized to elicit the collection of data for more comprehensive statistical analysis. However, the burden of proof that the SES intervention is effective should not be placed on the school districts. Rather, to promote the capacity for more extensive evaluation, firms that receive Title I funding from public school districts should be required to pass an independent third-party performance audit on an annual or semi-annual basis. The performance audit results, if based upon rigorous statistical analysis, would quantify the impact of tutoring and provide evidence of the effect of the tutoring program on aggregate student proficiency rates based upon established explanatory variables like those included and described in this study.

The idea of requiring SES providers to pass performance audits in order to be reauthorized to provide the tutoring services would likely incentivize tutoring firms to guarantee teacher credentialing standards and to collaborate more broadly with school or district officials so that tutoring is more individualized or targeted to supplement classroom curriculum and enrich participating students' learning opportunities. The compliance costs of annual or semi-annual performance audits, if mandatory as an eligibility requirement, may force smaller tutoring firms to leave the market or merge with other firms to consolidate program costs, which may effectively reduce the program variability related to the number of independent providers and may result in raising the overall quality of the tutoring programs.

Aside from requiring evidence of effectiveness, the NCLB provision should regulate the hourly rate that providers charge the districts for tutoring services. Such a reform may increase the Pareto improvement potential of the SES programs. The policy provisions could include a maximum rate that providers may charge the districts and establish a minimum number of tutoring sessions that they must provide to students who voluntarily enroll for the services. The price structure should maximize the number of tutoring hours or tutoring sessions in order to conform with the maximum per student expenditure that is established by the state according to the availability of Title I funds.

Local public education agencies often collaborate with private organizations to improve access to learning opportunities for all students. Policies that use public funds to develop such partnerships must also include empirical methods to evaluate the achievement of the policy objectives. In conducting the regression analysis for this study, I have shown that such an evaluation process is possible and practical at the local level with regard to the SES intervention component of the *No Child Left Behind Act* using the student-level data available through the Title I school districts. To replicate and improve the study, a researcher can expand the analysis

by including provider-level data. Requiring SES providers to exhibit a significant and positive effect on student test scores based on rigorous statistical analysis conducted by a third party would improve the efficacy of the policy without placing additional administrative burden on school districts that are mandated to provide the SES program for low-achieving students.

Appendix A

Research Methods, Variables, and Findings on Effect of Supplemental Educational Services

Authors and Publication Date	Methodology and Variables	Population and Sample Size	Results and Limitations
Muñoz, et al. (2008)	Quasi-experimental matching SES and control students based on chi-square tests for comparability on gender, ethnicity, poverty, and family structure, with ANOVAs for scores on assessment pretests. Linear regression using ANCOVA included final covariates of ethnicity, poverty, pre-test scores, and SES participation with year-end state assessment test scores as dependent variable	Urban district in KY with 30 PI schools; N = 1085: 4th, 7th, and 10th graders for reading outcomes; N = 916: 5th and 8th graders for math; 2005-06 school year	ANCOVA on year-end state test scores indicated SES program effect was not significant for reading (p=.20), math (p=.57); pretest scores was significant for reading and math (p<.001); ethnicity was significant for math (p<.001); poverty status was not a significant (p=.89) covariate for either Magnitude: Not sign. different from zero. Causal conclusions are restricted due to non-random sampling and inability to measure variables such as parental involvement, teacher effects, and SES attendance rates
Ross, et al. (2008)	Three models of multiple linear regression: (1) controlling for prior achievement using 2 yrs. scores on assessment tests, grade level, and SES provider; (2) as well as nesting teacher with grade level; or (3) nesting the matched-student pair (SES and non) by teacher and prior	Three districts in TN offering SES; sample selection criteria specified to achieve matched pairs of attendance among 21 SES providers: SES N = 610, math SES N = 470, reading Non N = 513, math Non N = 444, reading; 4th – 8th grades; 2005 2006 school year	Effect on 2006 test scores, SES provider not significant for math improvement and only for reading using Model 1 and 3; grade level significant for math using Models 1 and 3 and reading in each Model; prior achievement significant for improvement in math and reading using Models 1 and 2

Research Methods, Variables, and Findings on Effect of Supplemental Educational Services

Authors and Publication Date	Methodology and Variables	Population and Sample Size	Results and Limitations
Jones, et al. (2009)	Hierarchical Linear Modeling controlling for: individual student differences using gender, ethnicity, grade, poverty, performance level on state tests, IEP status, mobility, and school suspension; school differences using SES particip., other school program enrollment, and baseline scores on state assessment test	Chicago Public Schools (CPS) eligible for SES tutoring: N=272; disadvantaged students active in CPS during SES recruitment: N= 166,386 students; students registered for SES: N=46,856; participants in SES: N=37,095; 5^{th} – 8th grades; 2007-2008 school year	Controlling individual student differences indicated no overall significance of SES on state test scores ($p>.01$) for reading but significance for math ($p<.01$); small significance was indicated for students with below basic scores in math and reading. Study also followed SES providers, evaluating student retention, attendance, and gains on state test scores.
Heinrich (2009)	OLS regression model to estimate effect on state reading and math test scores over four years based on SES participation with any provider and with a particular provider; variables include student differences gender, ethnicity, poverty, English proficiency, IEP, GPA, absences and retention in school, foreign language class, prior year SES participation, grade, school attended, and propensity to register for SES	Milwaukee Public Middle and High School; report shows eval. of middle school test outcomes in 04-05: N=1562 for reading, N= 1571 for math; to eval. high school test outcomes in 04-05: N=1224 for math, N=1262 for reading; middle school test outcomes in 05-06: N=1075 for math N=1016 for reading; high school test outcomes for 05-06: N=2215 for math N=2200 for reading; data analyzed for school years 2003-04, 04-05, 05-06, and 06-07	Matched-pairs of SES or nonSES and baseline student and school variables indicated no significant difference in test scores; effects of total hours attended indicate small significance in 05-06 high school reading test scores ($p=.042$) only. Causal relationship restricted by possible unobserved variables, sample size, different standardized tests used during study, and different provider characteristics. Surveys and focus groups also used to evaluate/ qualify results.

Research Methods, Variables, and Findings on Effect of Supplemental Educational Services

Authors and Publication Date	Methodology and Variables	Population and Sample Size	Results and Limitations
Socias, et al. (2009)	Longitudinal data combined with student level fixed effects in regression model to evaluate effect of of SES participation on state standardized test scores using a z-score metric. Variables include SES eligibility, participation, provider, attendance, year, student tests scores	School district in Hillsborough Co., FL and Anchorage, AK permitted by USDE waiver to provided SES services to district students on a pilot basis in 2006-07 school year; grades 3-10 in AK with N=221 for math, N=258 for reading; and grades 2-12 in FL with N=1966 for math, N=7288 for reading	In AK, SES had no significant effect on test scores for math or reading in grades 3-10 from school years 2002 through 2007. In FL, SES participation had significant effect on test scores for math achievement only. Causal conclusion is limited by small size of AK sample; math achievement gains in FL SES participant scores is equal to .05 standard deviations more than non-SES student achievement scores in math
Barnett (2009)	OLS regression model to estimate CST test scores based on prior year's score and student-level effects including gender, ethnicity, English language classification, disabilities, GATE enrollment, and level of parent education; school-level effects included proportion of student demographics and district of attendance;	Los Angeles Unified School District; school level student populations of elementary, middle, and high school; N=246,299 SES eligible (54% of district population) 2007-08 school year	Comparing average MLR residuals, yields small SES effect on student achievement tests in English Lang. Arts (ELA) overall and across all three school levels, small effect on achievement in math overall, with out performance by SES participants at elementary and middle school levels; effect is equivalent to less than 5 scale pts. on achievement test

Research Methods, Variables, and Findings on Effect of Supplemental Educational Services

Authors and Publication Date	Methodology and Variables	Population and Sample Size	Results and Limitations
Barnett (2009), Continued	comparison of the residual (deviation from predicted and actual CST scores not explained by the OLS model) of SES and non-SES students for an estimated effect of SES participation on student achievement	(see above)	(see above)
Rickles (2007)	Linear regression to estimate 2005-06 achievement scores for SES eligible based upon 2004-05 scores and student variables of gender, English learner, disabilities, GATE, poverty, grade, and parent education level for eligible students; students with tests scores categorized by application, attendance, and frequency	Los Angeles Unified School District SES eligible students (N=216,192) who applied (23,086) for a program and actually attended (14,759) a program; students attending an ELA program: N=6538, a math program: N=8221;2005-06 school year	Controlling for student effects, SES had a small positive effect on test scores equal to .05 standard deviation or 2-3 scale score points for both math and ELA Causal conclusions restricted by unobserved variables such as student motivation, parent involvement, and SES provider effects

Appendix B

Interview Questions for Administrators[1]

MEASURING THE IMPACT OF NCLB MANDATED TUTORING

1. Effectiveness.
   - Does the district require and review pre- or post-test achievement data?
   - What assessment tools does the SES provider use to measure students' progress?
   - How does the SES provider meet the needs of ELL students and students with disabilities?
   - How does the SES tutoring increase student achievement in reading/language arts or math?
   - How much improvement and in what subject area have students made, to your knowledge and in your professional opinion?

2. Customer Satisfaction.
   - How does the district learn whether students who receive SES are satisfied with the tutoring?
   - How does the district learn whether parents of SES participants are satisfied with the tutoring?
   - Does the SES provider communicate effectively with principals, teachers, and parents about the progress of students enrolled in SES? If so, how often and by what method?
   - To your knowledge, what are teacher, student, and parent experiences with and reactions to the SES intervention, in general terms?
   - How does the district recruit enrollment to the SES program?

3. Service Delivery and Compliance.
   - What SES provider services does the district offer?
   - Do school administrators or faculty promote student participation in SES or have some type of influence on whether students or their parents chose the option to enroll in SES tutoring? If yes, what are the promotion efforts or what is the type of influence?
   - How does the SES provider comply with applicable state and district laws and contractual procedures associated with the delivery of SES?
   - How does the SES provider collaborate with principals, teachers, and parents to develop instructional plans designed to meet the needs of students?
   - How does the SES provider design instructional plans to align with state standards?

1

Ross, S. M., & Potter, A. (2006). *Evaluating Supplemental Educational Service Providers: Suggested Strategies for States.* U.S. Department of Education's Office of Elementary and Secondary Education (Award #S283B050057): Center for Research in Educational Policy, American Institute for Research, Academic Development Institute.

- How does the SES provider deliver the services it promised in its application or contract with regard to instructor credentials, instructional time, instructional plan, and communications with school, district, and parents about student progress?
- How often does the school, district, teacher communicate with SES providers?
- What amount does the district set aside from Title I funds for the SES intervention?
- How does the district fully expend the Title I funds set aside for SES intervention?
- Is the amount of Title I funds set aside adequate to provide SES for all interested students?
- Do administrators or faculty interact with SES providers to align tutoring activities with classroom curriculum?

References

Barnhart, M. K. (2009). *The Impact of Participation in Supplemental Educational Services (SES) on Student Achievement: 2007-08; Publication No. 2009-04.* Los Angeles, CA: Los Angeles Unified School District Research & Planning.

Bathon, J. M., & Spradlin, T. E. (2007, Fall). *Outcomes of the School Choice and Supplemental Educational Services Provisions of NCLB.* (C. f. Policy, Ed.) Retrieved October 1, 2011, from http://www.indiana.edu/~ceep/ http://ceep.indiana.edu/projects/PDF/PB_V5N8_Fall_2007_EPB.pdf.

Burch, P. (2007). Supplemental Education Services under NCLB: Emerging Evidence and Policy Issues. (Education Policy Research Unit, College of Education). Tempe, Arizona: Arizona State University. Retrieved from http://epsl.asu.edu/epru: http://epsl.asu.edu/epru/documents/EPSL-0705-232-EPRU.pdf.

Burch, P., Steinberg, M., & Donovan, J. (2007). Supplemental Educational Services and NCLB: Policy Assumptions, Market Practices, Emerging Issues. *Educational Evaluation and Policy Analysis*, 115-133.

California Department of Education. (2010-11). Retrieved February 15, 2012, from http://www.ed-data.k-12.ca.us.

California Department of Education. (2011, August). *2011 Adequate Yearly Progress Report Information Guide.* Retrieved September 2011, from http://www.cde.ca.gov/index.asp: http://www.cde.ca.gov/ta/ac/ay/.

California Department of Education. (2011, August 31). *Torlakson: Record Share of Schools Meet State Academic Goals.* Retrieved September 11, 2011, from http://www.cde.ca.gov/index.asp: http://www.cde.ca.gov/nr/ne/yr11/yr11rel62.asp.

California Department of Education. (2011, May). Retrieved October 2011, from http://www.cde.ca.gov: http://www.cde.ca.gov/ta/ac/ay/documents/overview11.pdf.

Center on Education Policy. (2006). *From the Capital to the Classrom: Year 4 of the No Child Left Behind Act.* Washington, DC: CEP.

Government Accountability Office. (2006). *No Child Left Behind: Education Actions Needed to Improve Local Implementation and State Evaluation of Supplemental Educational Services.* Washington, DC: GAO-06-758.

Heinrich, C. J., Meyer, R. H., & Whitten, G. W. (2010). Supplemental Education Services Under No Child Left Behind: Who Signs Up, and What Do They Gain? *Education Evaluation and Policy Analysis*, 32(2), 273-298.

Heistad, D. (2005). *Analysis of 2005 Supplemental Educational Services in Minneapolis Public Schools: An Application of Matched Sample Statistical Design.* Minneapolis, MN: Minneapolis Public Schools.

Jones, C. J. (2005). *SES Tutoring Programs: An Evaluation of the Second Year.* Chicago, IL: Chicago Public Schools Office of Research, Evaluation, and Accountability.

Jones, C. J. (2009). *The 2008 Supplemental Educational Services Program Year 5 Summative Evaluation.* Chicago: Chicago Public Schools, Office of Extended Learning Opportunities.

Minnici, A., & Bartley, A. P. (2007). *State Implementation of Supplemental Educational Services Under the No Child Left Behind Act.* Washington, DC: Center for Education Policy.

Munger, M. C. (2000). *Analyzing Policy Choices, Conflicts, and Practices.* New York City: W.W. Norton & Company, Inc.

Munoz, M. A., Potter, A. P., & Ross, S. M. (2008). Supplemental Educational Services as a Consequence of the NCLB Legislation: Evaluating its Impact on Student Achievement in a Large Urban District. *Journal of Education for Students Placed at Risk*, 13:1-25.

O'Connell, J., & Mitchell, T. (2008, June 20). Comments to the Notice of Proposed Rulemaking, Title I—Improving the Academic Achievement of the Disadvantaged, Docket ID: ED-2008-OESE-0003. Sacramento, CA. Retrieved October 2011, from http://www.cde.ca.gov: http://www.cde.ca.gov/nr/el/le/yr08ltr0620.asp#att.

Policy Analysis for California Education. (2008). *Conditions of California Education, 2008.* (D. Plank, Ed.) Retrieved October 2011, from http://gse.berkeley.edu/research/pace/reports: http://gse.berkeley.edu/research/pace/reports/Conditions_2008.pdf.

Potter, A., Ross, S., Paek, J., McKay, D., Ashton, J., & Sanders, W. (2007). *Supplemental Eduational Services in the State of Tennessee: 2005-2006 (2004-2005 Student Achievement Results).* Memphis, TN: University of Memphis Center for Research in Educational Policy.

Potter, A., Ross, S., Paek, J., Pribesh, S., & Nunery, J. (2006). *Supplemental Educational Services in the State of Louisiana: 2004-2005.* Memphis, TN: University of Memphis Center for Research in Educational Policy.

Rickles, J. H., & Barnhart, M. K. (2007). *The Impact of Supplemental Educational Services Participation on Student Achievement: 2005-06.* Los Angeles, CA: Los Angeles Unified School District Planning, Assessment and Research Division.

Rickles, J. H., & White, J. (2006). *The Impact of Supplemental Educational Services Participation on Student Achievement.* Los Angeles, CA: Los Angeles Unified School District Planning, Assessment, and Research Division.

Ross, S.M., & Potter, A. (2006). *Evaluating Supplemental Educational Service Providers: Suggested Strategies for States.* U.S. Department of Education's Office of Elementary and Secondary School Education (Award #S283B050057): Center for Research in Educational Policy, American Institute of Research, Academic Development Institute.

Ross, S. M., Potter, A., Pack, J., & McKay, D. (2008). Implementation and Outcomes of Supplemental Educational Services: The Tennessee State-Wide Evaluation Study. *Journal of Education for Students Placed at Risk*, 13:26-58.

School District Records. (2011-12). Retrieved from http://www.[district website].org/ses.

Socias, M., deSousa, J.-M., & Le Floch, K. C. (2009). *Supplemental Educational Services and Student Achievement in Waiver Districts: Anchorage and Hillsborough.* Washington, DC: U.S. Department of Education Office of Planning and Policy Development, American Institutes for Research, and RTI International.

Springer, M. G., Pepper, M. J., & Ghosh-Dastidar, B. (2009, August 12). Supplemental Educational Services and Student Test Score Gains: Evidence from a Large, Urban School District (Working Paper). Washington, DC: Urban Institute. Retrieved from http://www.urban.org/url.cfm?ID=1001308.

Studenmund, A.H. (2006). *Using Econometrics*. Occidental College: Pearson Education, Inc.

The White House, Office of the Press Secretary. (2011, September 23). *Obama Administration Sets High Bar for Flexibility from No Child Left Behind in Order to Advance Equity and Support Reform.* Retrieved October 6, 2011, from http://www.whitehouse.gov/briefing-room/statements-and-releases: http://www.whitehouse.gov/the-press-office/2011/09/23/obama-administration-sets-high-bar-flexibility-no-child-left-behind-.

U.S. Department of Education. (2004, September 14). *Choice and Supplemental Educational Services Frequently Asked Questions.* Retrieved October 2, 2011, from http://www.ed.gov: http://www2.ed.gov/parents/schools/choice/choice.html#11.

U.S. Department of Education: Office of Planning, Evaluation and Policy Development. (2009). *Implementation of the No Child Left Behind Act, Volume VII—Title I School Choice and Supplemental Educational Services: Final Report.* Retrieved October 2, 2011, from http://www.ed.gov: http://www2.ed.gov/rschstat/eval/choice/nclb-choice-ses-final.

Zimmer, R., Gill, B., Razquin, P., Booker, K., & Lockwood III, J. R. (2007). *State and Local Implementation of the No Child Left Behind Act: Volume I--Title I School Choice, Supplemental Educational Services, and Student Achievement.* Washington D.C.: U.S. Department of Education Office of Planning, Evaluation and Policy Development and the RAND Corporation.