NATIONAL CENTER ON
Performance Incentives

Supplemental Educational Services & Student Test Score Gains:

Evidence from a Large, Urban School District

Matthew G. Springer
Matthew J. Pepper
Bonnie Ghosh-Dastidar

Paper presented at the *NCLB: Emerging Findings Research Conference* at the Urban Institute, Washington, D.C. on August 12, 2009.

Conference Paper
August 2009

LED BY
VANDERBILT
PEABODY COLLEGE

IN COOPERATION WITH:
CALDER

The National Center on Performance Incentives (NCPI) is charged by the federal government with exercising leadership on performance incentives in education. Established in 2006 through a major research and development grant from the United States Department of Education's Institute of Education Sciences (IES), NCPI conducts scientific, comprehensive, and independent studies on the individual and institutional effects of performance incentives in education. A signature activity of the center is the conduct of two randomized field trials offering student achievement-related bonuses to teachers. The Center is committed to air and rigorous research in an effort to provide the field of education with reliable knowledge to guide policy and practice.

The Center is housed in the Learning Sciences Institute on the campus of Vanderbilt University's Peabody College. The Center's management under the Learning Sciences Institute, along with the National Center on School Choice, makes Vanderbilt the only higher education institution to house two federal research and development centers supported by the Institute of Education Services.

The conference was supported, in part, by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), funded by Grant R305A060018 to the Urban Institute from the Institute of Education Sciences, U.S. Department of Education and the National Center for Performance Incentives (NCPI) at Vanderbilt University. This study was supported, in part, by the National Center on School Choice at Vanderbilt University, which is funded by the United States Department of Education's Institute of Education Sciences, Grant R305A040043. The authors appreciate helpful comments and suggestions from Dale Ballou, Mark Berends, J.R. Lockwood, and Ron Zimmer and seminar participants at the annual meetings of the American Education Finance Association and the Association of *Public Policy Analysis and Management*. They also acknowledge the many individuals at the school district for providing data and expert insight to conduct their analyses. The views expressed in the paper are solely those of the authors and may not reflect those of the funders or supporting organizations. Any errors are attributable to the authors.

Please visit **www.performanceincentives.org** to learn more about our program of research and recent publications.

Supplemental Educational Services and Student Test Score Gains: Evidence from a Large, Urban School District

MATTHEW G. SPRINGER
Vanderbilt University

MATTHEW J. PEPPER
Vanderbilt University

BONNIE GHOSH-DASTIDAR
RAND Corporation

Abstract

This study examines the effect of SES on student test score gains and whether particular subgroups of students benefit more from NCLB tutoring services. The sample used includes information on students enrolled in 3rd through 8th grades nested in 121 elementary and middle schools over a five-year period from 2003-04 to 2007-08. A total of 17 elementary and middle schools were required to offer SES at some point during the period under study, and 9,861 student-year pairings in the sample were eligible to receive SES. While the authors' preferred analytic approach takes advantage of panel data to implement an analysis strategy with student fixed effect regression methods, they also test the robustness of these estimates to a number of alternative approaches, including a comparison of student test score gains between current and future SES participants. The authors find consistently significant and positive average effects of SES on test score gains in mathematics. Results in reading tend to be insignificant. SES tutoring does not appear to disproportionately benefit a particular racial/ethnic group or ability level. Female students and students with disabilities appear to benefit more from participating in SES. SES has a significant, cumulative effect on students in both mathematics and reading. They also demonstrate that not accounting for content area of tutoring can cause downward bias in estimates of the SES treatment effect. These findings are qualified on a couple of dimensions.

1. Introduction

After-school programs have become a popular approach to enhance academic opportunities and outcomes of public elementary and secondary school children in the United States. Advocated during the 1980s as a strategy for lowering juvenile crime, increasing social and job-readiness development, and lessening opportunity costs for single parents to enter the workforce (Catalano et al, 1999; Gelbach, 2002; Connely, 1992; Blau and Robins, 1998), the purpose of after-school policies became increasingly academically-oriented in the 1990s when federal legislation established the 21st Century Community Learning Centers program.[1] The supplemental educational services (SES)[2] provision of the No Child Left Behind Act of 2001 (NCLB) further reinforced the intent of after-school programs to "increase the academic achievement of eligible children on academic assessments…and attain proficiency in meeting the State's academic achievement standards" (Title I, Section 1116(e)(12)(C)).

SES are free tutoring services offered to low income children in low performing schools outside of the hours of the typical school day. SES are provided through a variety of entities including for-profit, non-profit, local community, school district, and college and university tutoring programs. Students who both are identified as low-income and attend a Title I school that has failed to make adequate yearly progress for three or more consecutive years under a state's NCLB accountability policy can enroll with a service provider.[3] With an approximate $2.78 billion included

[1] In 1994, Congress authorized the 21st Century Community Learning Center to open schools for broader use in their communities. The program became more narrowly focused on school-based academic and enrichment activities in 1998, growing from an appropriation of $40 million per year in 1998 to more than $1 billion in 2002 (Dynarski et al, 2003).

[2] In this paper, the acronym SES refers to supplemental education services which should not be confused with socioeconomic status.

[3] The U.S. Department of Education has entered into flexibility agreements with five districts and eleven states for the 08-09 school year to offer SES to eligible students after only two years of failing AYP.

for NCLB's SES provision in the President's FY 2008 budget request to Congress (representing a 58.7 percent funding increase since first being funded in 2001), the federal government is placing significant weight on after-school programs to improve academic opportunities and outcomes in public elementary and secondary schools.

However, the evidence paints a mixed picture of the effects of pre-NCLB after-school tutoring programs on student test scores. Several experimental and quasi-experimental evaluations report inconclusive or insignificant effects (Reisner et al, 2002; Walker et al, 2003; Dynarski et al, 2003). A meta-analysis of out of school time that does not include those studies indicates a small but statistically significant and positive effect on mathematics and reading test scores (Lauer et al, 2006). Yet, two surveys of the literature call into question the credibility of the evidence on the benefits of after-school tutoring (Hollister, 2003; Kane, 2004).

The literature assessing the effects of SES on student achievement is similarly inconclusive. Evaluations have found positive program effects in mathematics and reading (Chicago Public Schools, 2005; Rickles and Barnhart, 2007; Zimmer et al, 2006; Zimmer et al, 2007), while other studies report mixed (Chicago Public Schools, 2007; Heistad, 2007; Rickles and White, 2006) or negligible program effects (Potter et al, 2007; Heinrich, Meyer, and Whitten, 2007). We do not know of any studies, to date, concluding that SES had a negative effect on student test scores.

The SES evaluation literature also varies widely with respect to methodological rigor. None of the existing evaluation studies employ a random assignment design,[4] which makes selection bias a salient concern because schools are required to offer SES for repeatedly failing to make AYP and student participation in SES is voluntary. However, even though there are statistical and

---

Locations include Boston Public Schools, Chicago Public Schools, and the states of Arkansas, Florida, and Utah.

[4] Two randomized, controlled trials compare the impact of students participating in a typical after-school program to those students participating in an after-school tutoring program with a specific curriculum (Fitzgerald and Hartry, 2008; Black et al, 2008).

econometric methods aimed at attenuating the influence of selectivity bias when evaluating a policy intervention in the absence of randomization, the majority of SES evaluations do not implement a rigorous, non-experimental design. Indeed, in our review of the literature, we identified only four studies that did so (Zimmer et al, 2006; Zimmer et al, 2007; Heinrich, Meyer, and Whitten, 2007; Heistad, 2007).

Other forms of omitted variable bias may skew estimates of the SES parameter measured in previous evaluations. For example, most studies do not account for differential patterns of student enrollment and attendance in SES. Moreover, no study distinguishes the content area in which a student receives tutoring, even though a student may receive tutoring in reading, mathematics, or both subjects. Both the failure to show up by students enrolled with a provider and the variation in the content area in which students receive tutoring may dilute estimates of the SES treatment effect.

In an effort to more precisely identify the effects of SES on student test scores, and whether particular subgroups of students benefitted more from tutoring services, we constructed a comprehensive panel data set in partnership with a large, urban school district in the southern United States. The district enrolls over 70,000 students, 72 percent of who qualify for free or reduced price lunch. Our sample includes students in 3rd through 8th grades nested in 121 elementary and middle schools over a five year period comprising the 2003-04 to 2007-08 school years. Seventeen schools were required to offer SES, and 9,861 student-year pairings were eligible to receive SES. In total, there are 114,978 student-year observations.

Our basic modeling strategy relies on a student fixed effects model to control for unobserved, time invariant sources of heterogeneity between SES and non-SES students. We estimate how a student who attends SES performs compared to how that student is expected to have performed without SES tutoring. In addition to the usual student- and school-level covariates,

we account for actual attendance of students enrolled in SES (after enrollment), the content focus in which a student receives tutoring, and/or the number of years a student attended SES.

To test the robustness of our main findings, we adopt two cross-sectional methods for estimating the SES treatment effect. First, we implement a modeling strategy following the framework employed in Zimmer et al.'s (2007) assessment of NCLB's student transfer provision. We compare the performance of students enrolled in SES to future SES participants, where future participants are defined by those students who were not yet eligible for SES but elected to enroll with a provider when they became eligible in the following school year. The approach accounts for selection bias in that the comparison group, by its future participation, has signaled a willingness of enrolling in SES if it were available to them. Admittedly, if student performance in the year prior to enrollment is correlated with enrolling in SES, comparisons of current and future participants will produce biased estimates.

A second cross-sectional method uses propensity score analysis as first defined by Rosenbaum and Rubin (1983) and advanced by Lunceford and Davidian (2004). Propensity score analysis balances nonequivalent groups in an effort to more reliably estimate the effect of a policy intervention using observational data. However, this approach may be suboptimal in the current context.[5] Most education data sets contain limited number of covariates about students and their situation, and factors not observed in the data that are associated with student test scores and selection into tutoring services are likely to confound estimates of the SES treatment effect.

[5] Researchers have examined whether non-experimental methods replicate estimates on the experimental impact of a program or policy intervention. Agodini and Dynarski (2004) concluded that propensity score analysis performed poorly when measuring the experimental impacts of 16 dropout prevention programs. Wilde and Hollister (2007) reached a similar conclusion using data from the Tennessee STAR study. Diaz and Handa (2006: 341) suggest that propensity score analysis can adequately address selection bias if researchers have "…an extremely rich set of covariates, detailed knowledge of the beneficiary selection process, and the outcomes of interest [are] measured as comparably as possible."

However, as we have several years of pre-intervention data, we do incorporate pre-SES test scores, which may capture some of the unobserved correlates of student achievement.

In this study, we find significant and positive effects of SES on student test scores in mathematics. Results in reading tend to be inconsistent. Findings are maintained when controlling for either the percent of or absolute hours of SES tutoring sessions attended. The magnitude of the SES treatment effect also increases in the expected direction when controlling for the content area of tutoring. Furthermore, the results from the student fixed effect models align with comparisons of SES students with the pre-enrollment gains students who reveal an SES enrollment preference through future participation.

SES is measured to have a significant cumulative impact on test score gains in both mathematics and reading if a student receives two or more years of tutoring. SES tutoring does not appear to disproportionately benefit a particular ethnic group or ability level, where ability level is determined by the previous year's quartile performance. However, female students and students with disabilities disproportionately benefit from participating in SES.

Our findings are qualified on several dimensions. First, observational studies of this nature are always susceptible to selectivity bias or other spurious relationships arising from events or processes such as differential teacher effects. Second, we do not know the extent to which a student actually received academically focused tutoring when he attended SES, nor do we know the quality of those services.[6] Third, even though most SES providers serve clients in urban school systems, it is unclear if students enrolled in tutoring services and the services offered in the district under study can generalize to those of other urban school districts.

The subsequent paper is divided into six sections. In section 2, we provide a brief overview of the status and trends of SES in the United States and the school district studied in this paper. In

[6] Provider-specific models were run by the authors. While variation did exist in the efficacy of provider, it was not the case that a small subset of providers drove the results.

Section 3, we review relevant literature on the impact of SES on student test score gains, paying particular attention to shortcomings in prior evaluations of NCLB's SES provision. Sections 4 and 5 describe our analytic strategy and our data and sample, respectively. Findings are presented in Section 6. Section 7 discusses results and implications for SES policy.

2. No Child Left Behind and Supplemental Education Services

The SES provision is one of several regulation found in NCLB, including student transfers from failing schools, school restructuring, and state-level takeovers. SES are offered through a variety of entities, including for-profit, non-profit, local community, local education agencies, and college and university tutoring programs. A national survey on the incidence of SES finds that 54 percent of SES providers are private for-profit companies, 21 percent are nonreligious non-profit entities, and nine percent of SES providers are LEA operated (Center on Education Policy, 2006). The remaining 16 percent of SES providers are operated by either local community groups or colleges and universities.

A recent review of the literature reported that approximately 13 percent of Title I schools in the United State were required to offer SES during the 2007-08 school year, where between zero and approximately 70 percent of eligible students enrolled (Springer et al., 2009). Although the number of districts with schools required to offer SES has remained fairly stable over time, the proportion of schools within these districts increased nearly 2.5 times (from 27 percent to 65 percent) from the 2002-03 to 2005-06 school years (Center for Education Policy, 2006). Studying the effects of SES on student test scores in urban school districts is particularly important bearing in mind that less than one third of Title I schools are located in urban settings, yet more than half of all urban Title I schools have failed to make AYP for three or more consecutive years (Center for Education Policy, 2006).

Table 1 displays descriptive information on the status and trends of the SES marketplace in the district under study. A total of 20 SES providers offered tutoring services to eligible students at some time during the 2003-04 to 2007-08 school years. The provider with the largest market share of students enrolled 20.5 percent of all students enrolled in SES when pooled across years, while the next five largest providers accounted for another 58.5 percent of enrolled students. The market share for any single provider remaining was fewer than 5 percent of students enrolled in SES. Perhaps strikingly, 60 percent of the 20 providers stopped offering tutoring services in the school district after one school year.

[Insert Table 1 Here]

The total number of Title I schools required to offer SES steadily increased from the time the first school in the district was required to offer SES (i.e., the 2003-04 school year). An additional five schools were required to offer tutoring services in the 2005-06 school year, and by the 2007-08 school year, a total of 14 schools had to offer SES (about 12 percent of those schools in the district with at least one grade in the 3rd through 8th grade range). The total number of students eligible to receive SES grew from approximately 100 students in the 2004-05 school year to more than 4,300 students in the 2007-08 school year, or about 9,900 students across all years.

The percentage of students enrolled in SES by year ranged between 18.4 and 23.0, with an average enrollment rate of 21.0 percent, which is similar to national estimates published by the Government Accountability Office (Shaul, 2006). However, the percentage of students that actually received tutoring services fell between 26.4 and 35 percent after taking into consideration those students that never attended a tutoring session.

Figure 1 displays the SES take-up rate among students enrolled in elementary schools and middle schools from 2005-06 through 2007-08 school years. Each vertical bar represents the percent of a school's eligible student population that attended at least one hour of SES in a

particular school year. Panel A reports these SES take-up rates for elementary schools and Panel B does the same for middle schools. Not all schools have observations for each of the three school years because eligibility is determined by AYP status, which varies from year to year.

[Insert Figure 1 Here]

Two trends are particularly noticeable in Figure 1. First, elementary schools have much greater success at enrolling students for SES. More than one-third (39 percent) of eligible elementary school students signed-up for SES, and 87 percent of those students attended at least one tutoring session. In contrast, 19 percent of middle school students eligible for SES signed-up. Only seventy-two percent of those middle school students attended at least one tutoring session.

The data displayed in Figure 1 further demonstrates that SES enrollment rates vary from school to school and, to a lesser extent, within the same school from one school year to the next. Anecdotal information provided by administrators operating the district's Office of Federal Programs pointed out that a motivated school administrator or site-based SES coordinator can significantly affect the SES take-up rate. Furthermore, district- and school-level personnel have become more familiar with their roles and responsibilities and worked to improve their delivery of SES over time.

3. Review of Relevant Literature

Very few studies have attempted to estimate the impact of SES on student outcomes. None of these studies use a random assignment evaluation design due, in part, to NCLB legislation being applied to all schools and school systems when the law was first enacted in 2002. Even though statisticians and econometricians have developed several strategies for estimating the effect of a policy intervention in the absence of randomization, only a handful of studies have implemented

rigorous non-experimental methods (e.g., statistical matching techniques or panel data methods) to draw inferences about the SES treatment effect.[7]

Two studies attempt to correct for selection bias using propensity score analysis, both of which reported negligible effects of SES on average test score gains in mathematics and reading (Heinrich, Meyer, and Whitten, 2007; Heistad, 2005). Propensity score analysis assumes the conditional probability of treatment and comparison condition students as a function of observable characteristics of students and schools can be used to balance the covariates among SES participants and non-participants. Making use of propensity scores to estimate the SES treatment effect is likely to produce biased conclusions because factors associated with student test scores and likely influence student selection into SES are typically unobserved in most education databases (i.e., parent motivation or innate student ability). Indeed, most data sets available to education researchers have a limited amount of information about students, their teacher and schools, and their home environment.

Researchers frequently use panel data to implement fixed effect regression methods that can account for latent heterogeneity among students, teachers, classrooms, and so on. So long as unobserved characteristics do not change overtime, fixed effect methods can produce reasonable inferences about an educational intervention where assignment of units is outside the control of the researcher. If a limited amount of information is available about the units under study, fixed effect methods offer several advantages over propensity score analysis. However, they require a minimum of three years of valid test score information per student to get two gain scores (one pre- and one-post treatment) and can result in biased estimates if the subsample of students used to identify the treatment effect is not representative of all students.

---

[7] Gill and Deke (2008) recently designed a large-scale evaluation of SES for the United States Department of Education's Institute of Education Sciences, which takes advantage of SES oversubscription to compare accepted and denied applicants using a regression discontinuity framework. See Springer et al. (2009) for a more complete review of the literature.

Contrary to the estimates reported in the two SES evaluations that relied on propensity score analysis, studies that employed a student-fixed effect approach report generally positive and significant effects on student test scores. For example, Zimmer et al.'s (2006) analysis of two after-school tutoring programs in Pittsburgh reported average effect sizes of 0.26 in mathematics and negligible effects in reading. Zimmer et al (2007) report that, on average, participation in SES increases test score gains in mathematics by 0.09 standard deviations and 0.08 standard deviations in reading. Furthermore, students enrolled for two or more years was even larger, equivalent to 0.17 and 0.15 standard deviations in mathematics and reading, respectively.

Most studies reporting a SES treatment effect on student test scores are also susceptible to bias from incomplete information on patterns of student enrollment and attendance. For example, even though three of the four rigorous, non-experimental evaluations tracked student attendance patterns, only two of those studies had adequate data to calculate a student's attendance rate at SES. Attendance rate can proxy for a student's exposure to tutoring, but may hide large differences in the number of hours provided, which vary by provider. At the same time, some empirical work contends the amount of time spent in after-school tutoring may be less important than what occurs during tutoring sessions or whether that time was actually spent on academics (Aaronson, Zimmerman, and Carlos, 2002; Karweit, 1985; Lauer et al., 2006).

Finally, studies evaluating the effect of SES on student achievement have not accounted for the content area in which a student receives tutoring. Not knowing the subject(s) in which a student receives tutoring may cause the measured impact of SES on mathematics (or reading) test score gains to be biased downward when using the entire SES population because students who received tutoring in reading (or mathematics) only would dilute the treatment effect (assuming, of course, there are negligible spillover effects across subjects). Omitting content area of tutoring is very salient in the context of the current study considering 21.8 percent of students enrolled in SES

received tutoring in mathematics only, 32.1 percent received tutoring in reading only, and 46.1 percent received tutoring in both mathematics and reading.

In summary, our review of relevant literature highlights several limitations in the existing knowledge base. First, none of the existing evaluations use randomized, controlled trial designs which makes it difficult to answer how a student who participated in SES would have fared without the program. Second, in a majority of studies, researchers have attempted to estimate the effect of SES on student test scores without taking into account selection bias and, as a result, conclusions drawn from these studies are likely wrong. Third, of the four rigorous evaluations reviewed, two rely on statistical matching techniques to correct for selection effects. Since estimates of program effects tend to be highly sensitive to the vector of variables used to create propensity scores, and we know the amount of information available in most education databases tends to be limited, it is questionable whether these studies provide reliable estimates. Finally, most studies are likely susceptible to omitted variable bias because their analytic strategy omit patterns of student enrollment and attendance in SES as well as the content area in which a student receives tutoring.

4. Analytic Strategy

4.1. *Longitudinal Analysis with Student Fixed-Effect Approach*

Our base model for estimating the relationship between SES tutoring and student test score gains can be expressed as:

$$\Delta Y_{ijt} = Y_{ijt} - Y_{ij,t-1} = \propto_0 + \propto_1 (registered)_{it} + \propto_2 (student)_{ijt}$$
$$+ \propto_3 (school)_{jt} + \gamma_i + \theta_{gt} + v_{ijt} \quad (1)$$

where, ΔY_{ijt} is the spring-to-spring standardized test score gain in reading or mathematics for student *i* attending school *j* in year *t*; $registered_{it}$ is an indicator variable that takes a value of one

if student *i* signed-up for SES in year *t* or zero if a student does not register for SES; $student_{it}$ is a vector of observable student-level characteristics for student *i* in year *t*; $school_{jt}$ is a vector of year-specific school characteristics; γ_i is a student fixed effect; θ_{gt} is a year by grade effect which controls for changes in the test, changes in how well aligned the test is with curricula, and student cohort effects; and v_{ijt} is the random disturbance term.

The parameter of interest in Equation (1) is the coefficient of $registered$, α_1, which represents the average SES treatment effect on student test score gains in mathematics or reading. The parameter α_1 reflects potential test score gains from offering free tutoring services to low income children in low performing schools. However, $registered$ is likely an imperfect measure for whether or not a student benefits from tutoring because α_1 does not differentiate between students registered for SES and those students registered for SES that attended tutoring services; nor does Equation (1) account for the content area in which a student receives tutoring.

Consequently, we substitute $registered$ with the indicator variable, $attended$, which takes a value of one if student *i* attended at least one tutoring session in year *t* or zero if a student does not register for SES, registers but does not attend a single tutoring session, or is ineligible for tutoring services. The coefficient of $attended$ represents the average difference in mathematics or reading test score gains between students that attended at least one tutoring session and students that never attended a tutoring session. Select specifications, as reported in the next section, also account for the content area in which a student receives tutoring.

Our base estimation strategy accounts for potentially confounding factors associated with latent characteristics of a student (i.e., motivation, family characteristics, or parental inputs) through replacement of an additive individual effect with a student fixed effect, γ_i. A student fixed effect approach may produce biased estimates if the subsample of students used to identify the SES

treatment effect is not representative of all SES students in the sample, or the necessary assumption of unobserved characteristics being time invariant is invalid. For example, the complex information networks among the poor may result in an endogenous relationship between the likelihood of signing-up for SES and the exposure of a student's peers to SES; that is, a student's likelihood of signing-up for SES increases as he (or his parents) encounters other peers who attend tutoring services.

We also examine the potential test score differences among those students that chose to attend SES and those that did not register or registered and did not attend. We use a modified form of Equation (1), which can be expressed as:

$$\Delta Y_{ijt} = Y_{ijt} - Y_{ij,t-1} = \alpha_0^{'} + \alpha_1^{'}(registered)_{it} + \alpha_2^{'}(attended)_{it} + \propto_3^{'} (student)_{ijt}$$
$$+ \propto_4^{'} (school)_{jt} + \gamma_i + \theta_{gt} + v_{ijt} \qquad (2)$$

where, $\alpha_0^{'}$ is the average difference in mathematics or reading test score gains between students that registered for SES and students that never registered for SES and $\alpha_2^{'}$ is the average difference in mathematics and reading test score gains between students that registered for SES and students that registered for SES and attended at least one tutoring session. We are most interested in the estimate on $\alpha_0^{'} + \alpha_1^{'} + \alpha_2^{'}$, which reflects the average test score gain for students that registered for SES and attended at least one tutoring session. The estimate on $\alpha_2^{'}$ is also of interest in that $\alpha_2^{'}$ differentiates the average effect of SES registration and attendance on test score gains in mathematics and reading.

We further explore this line of inquiry through the inclusion of exploratory variables that broaden the model specification identified in Equation (2). We first introduce a continuous control variable for hours of attendance. The continuous attendance variable is included as either the percent of total hours attended or the absolute hours attended, which vary by student and within provider. Although these variables arguably offer an adequate proxy for student exposure to

tutoring services, previous research literature contends the amount of time spent in after-school tutoring may be less important than what occurs during tutoring sessions or whether that time was actually spent on academic learning (Aaronson, Zimmerman, and Carlos, 2002; Karweit, 1985; Lauer et al., 2006).

We also create three binary variables indicating if a student received tutoring in mathematics only, reading only, or both mathematics and reading. We do not anticipate spillover effects between subjects; that is, a student who receives tutoring in reading only (or mathematics only) is not expected to perform better in mathematics (or reading) as a result of tutoring services. Finally, we investigate the relationship between SES treatment and observable student characteristics (i.e., race/ethnicity, gender, limited English proficient (LEP) students, students eligible for special education services, prior ability level). We also investigate the benefits of attending SES one year compared to attending more than one year.

4.2. *Current and Future Participants Estimation Strategy*

We implement a second estimation strategy that compares test score gains of students enrolled in SES to future SES participants. Comparing the test score gains of a matched sample of students currently enrolled in SES to test score gains of students who were not yet eligible for SES but elected to enroll with a provider when they became eligible in the following year is another way to control for unobservable factors under the assumption students have equal propensity to enroll when given the opportunity. Modeled after Zimmer et al. (2007), this estimation strategy can be expressed as:

$$\Delta Y_{ijt} = Y_{ijt} - Y_{ij,t-1} = \delta_0 + \delta_1(attended)_{it} + \delta_2(student)_{ijt}$$
$$+\delta_3(school)_{jt} + \theta_{gt} + v_{ijt} \qquad (3)$$

where, all coefficients are similar to those reported in Equation (1) and the parameter of interest is the coefficient of $attended$, δ_1, which represents the average SES treatment effect holding constant all observed student and school attributes. Equation (3) will produce biased results of the SES treatment effect if the timing of participation in SES were related to a student's test score in the year immediately preceding enrolling in SES or the sample of current and future participants is not representative of all SES participants. Because the sample is restricted to 714 control and 1,020 treatment observations, we only use the future participant identification strategy to estimate the average SES effect.

4.3. *Propensity Score Analysis*

Our final estimation strategy employs propensity score analysis. Propensity score analysis is a statistical technique implemented to balance two non-equivalent groups on observed covariates to get more accurate estimates of the average treatment effect under the assumption that bias will be reduced when the comparison of students enrolled and not enrolled in treatment are as alike as possible. Propensity scores summarize all the information from the observed covariates into one single number, namely, the probability of being assigned to the treatment given by the covariates. Selection bias is eliminated only if the exposure to treatment can be considered to be purely random among individuals who have the same value of the propensity score; that is, no unobserved covariates are confounded with treatment within the same levels of the propensity score.

We specified a student's propensity of enrolling in and attending SES using the following logit model:

$$\widehat{S_i} = P(D_{ijt} = 1|C_i) \quad (4)$$

where, $\hat{S}_i$ is student *i*'s propensity of enrolling in and attending SES during the window of the study; D_{ijt} is a binary variable taking on a value of one if student *i* in school *j* enrolled in SES or zero if a

student did not enroll; C_i is a vector of observed student-level pre-treatment covariates including mathematics and reading test scores, race, gender, eligibility for free lunch, eligibility for reduced price lunch program, LEP status, special education status, student attendance rate at school, and grade-level of student.

After fitting the logit model and estimating propensity scores, we divided the propensity scores into quintiles. To check for distribution of the quintiles across the SES and non-SES students, we then ran a series of analysis of variance models with each student-level covariate used in the adjustment as the dependent variable, and propensity score quintile, SES treatment status, and the interaction of quintile with SES as independent variables. We also tested for balance among the covariates by conducting a t-test for the two groups prior to any adjustment using the propensity score.

We incorporated the propensity score into our evaluation models in two ways. In the first direct way, we estimate the weight as the inverse of the propensity score for the treatment group, $w_i = 1/\hat{S}_i$. For the control group, the weight equals $w_i = \frac{1}{1-\hat{S}_i}$. We include these weights when estimating the following regression model:

$$\Delta Y_{ijt} = Y_{ijt} - Y_{ij,t-1} = \beta_1(attended)_{it} + \beta_2(student)_{ijt} + \beta_3(school)_{jt} + \theta_{gt} + v_{ijt} \quad (5)$$

Thus, the average SES effect is a weighted average of the outcome or gain scores can be expressed as:

$$Ave\left(Y_{ij,t} - Y_{ij,t-1}\right) = \frac{\sum w_{ij} \times (Y_{ijt} - Y_{ij,t-1})}{\sum w_{ij}} \quad (6)$$

In an alternate approach, we used the propensity scores to adjust for covariates through stratification, and then the strata variable is included as a covariate when calculating the effect of SES. The quintiles of the distribution of the propensity scores form five strata, thus effectively

making a coarse match between SES and non-SES students. There were sufficient SES and non-SES kids in three of the five strata, while in the two lowest strata contained an insufficient number of SES students to estimate the SES effect within each stratum.[8] We created a series of indicators for the strata and included these as covariates in the model where we estimate the SES effect.

5. Data and Sample

We cleaned and merged relevant student, school, and provider information from multiple data sources to create a single longitudinal data file for a five-year period comprising the 2003-04 to 2007-08 school years. Data were drawn from management information systems maintained by the school district, including test score files, enrollment history files, and federal program files. In total, our sample includes approximately 143,801 continuously enrolled student-year observations in mathematics and reading nested in 121 elementary and middle schools.[9]

The test score file is a flat file that contains annual test score results for mathematics, reading, science and social studies. We focus on data from the mathematics and reading assessments because test scores are linked across grades and presented on a single developmental scale. A one to two percent error rate in the unique identifier linking a student to his test score was corrected on a case-by-case basis, resulting in student-score match rates of greater than 99.9 percent across all years and grades.

We standardized student test scores by subject, grade, and year and then constructed a simply standardized gain score by subtracting scores at time *t* from those at time *t-1*. A gain score indicates a student's test score is below the mean for all tested students in that subject, grade, and

[8] The two lowest strata contained only two student observations that enrolled in SES.

[9] "Continuously enrolled" is defined using the state's definition under NCLB; that is, a student must be enrolled within the same school from the twentieth day of school through mid-April when the state's high-stakes assessments are administered.

school year, while a positive score indicates a student's test score is above the distribution mean. A standardized gain score of zero means a student test score from one year to the next increased the average amount for that grade, year, and subject.

The enrollment history file contains student demographic information such as a unique student identifier, race, gender, date of birth, grade, free lunch status, and reduced lunch status. The file also provides a transactional enrollment history which records dates of school enrollment and transfer for each student. The enrollment history file was supplemented with daily student attendance records to create an in-school attendance variable for each student.

The federal program file tracked the involvement of an individual student in SES on several dimensions, including student enrollment, total hours scheduled, total hours attended, the name of the tutoring provider, and the content area of tutoring (i.e., mathematics, reading, or both). Under mandate by the state department of education, this data is recorded and maintained by a designated SES coordinator at the district. SES attendance information is tracked through invoices submitted by providers. School-level SES coordinators confirm the accuracy of records in the federal program file at regular intervals throughout the school year.

Table 2 displays summary statistics on select characteristics of students and schools. Information is reported by all students in the district, students in schools required to offer SES, students eligible to receive SES in failing schools, students who signed-up for SES, students who attended at least one SES tutoring session, and students who enrolled but did not attend SES. The final set of columns report attendance by SES content area for students who attended at least one tutoring session.

[Insert Tables 2 Here]

Roughly half of the student observations in the sample are female, 47.8 percent are Black, 36.4 percent are white, and 12.4 percent are Hispanic. Approximately 62 percent of students in

grades 3 to 8 qualified for free or reduced price lunch program. The great majority of those students are part of the free lunch program (≈ 87 percent). More than 10 percent of SES eligible students received special education services with the largest share of those students receiving between 5 and 20 hours of service per week. The average daily in-school attendance rate was 95.7 percent.

On select demographic characteristic, students attending a school required to offer SES are noticeably different from the average student in the district. A smaller percentage of white students enrolled in failing schools, which is offset by failing schools enrolling a larger concentration of Hispanic and Black students. Approximately 81 percent of students in failing schools qualify for the free- or reduced-price lunch opposed to an average of 62 percent of students in the district. There are modest differences in the percentage of students receiving more than 20 hours of special education services, while virtually equal shares of students qualify for Title I lunch program in schools offering SES and those not offering SES. Students classified as limited English proficiency (LEP) are much more likely to attend a school required to offer SES.

Table 2 further indicates students eligible for SES are different from the average student in the district and the average student in schools offering SES. A slightly larger percentage of eligible students are Black and Hispanic, while fewer white students are eligible for SES. All of these students are eligible for free or reduced price lunch. Students eligible for SES score, on average, 6 to 17 points lower than the average student in the district on standardized assessments in mathematics and English language arts.[10] This is equivalent to between 0.17 and 0.43 standard deviations dependent upon the grade and subject under consideration.

Students enrolled in SES are different from the average student eligible to receive SES on a number of observable characteristics. A greater percentage of Black students enroll, and enroll and attend SES, while a lower percentage of eligible, white students enroll. Hispanic students enrolled in

[10] Means and standard deviations of level scores and gain scores are reported by grade and student subgroups in Appendix A.

SES are slightly less likely to attend SES sessions. In addition, special education students receiving between five and 20 hours of special education services per week are more likely to take-up SES. Female students are slightly more likely to receive tutoring in mathematics only, while Hispanic and LEP students are more likely to receive tutoring in reading only. Students receiving between 5 and 20 hours of special education services each week disproportionately receive SES tutoring in mathematics.

Table 2 also reveals that, as students matriculate into middle school, subject-specific tutoring slightly shifts from reading only to mathematics only. An investigation into subject area data by SES providers reveals that providers generally offered services either focused on both subjects or focused on one subject only. For example, the provider that tutored the most students focused only on mathematics for 133 students, only on reading for 151 students, and both or unknown for 8 students. The third largest provider focused on both subjects for 199 students, and a single subject only for 15 students.

6. Results

6.1. *The Average Effect of SES on Student Test Score Gains*

Table 3 displays results for the estimated effect of SES on mathematics and reading test score gains. Panel A and Panel B report results when mathematics and reading test score gains are the dependent variable, respectively. All model specification include grade-by-year and student fixed effects, while estimates in the second column of each panel adds controls for characteristics of students. Our preferred model specification, as displayed in the third column of each panel, contains student fixed effects, grade-by-year effects and time-varying student and school characteristics.

[Insert Table 3 Here]

Models (1) through (6) show a positive, statistically significant average effect of SES on student test score gains in mathematics and reading. Students registered in SES experienced increases in test score gains of 0.088 standard deviations in mathematics and 0.076 standard deviations in reading. However, using students identified as enrolled in SES is an imperfect strategy to measure whether a student benefits from SES because students that registered for SES and those students that registered for SES and attended tutoring services are considered the same.

As displayed in the bottom half of Table 3, Models (7) through (12) estimate the average difference in mathematics or reading test score gains between students that attended at least one tutoring session and students that never attended a tutoring session. The precision and magnitude of the estimates are virtually identical when test score gains in mathematics provided the outcome measure. Estimates from Models (10) through (12) further suggest that students that enrolled in SES but did not attend a single tutoring session appear to be driving the positive SES treatment effect in reading. The magnitude of the SES treatment effect is about half as large as models fitted using registered and enrolled students, and the estimates for the SES treatment effect are no longer significant at conventional levels. Subsequent tests do not reveal statistically meaningful test score gain differences between students that enrolled for SES and those students that enrolled and attended at least one tutoring session in either mathematics or reading.

A number of dynamics can help to explain these counterintuitive results. First, a comparison of test scores in both the year prior to and the two years prior to a student enrolling in SES reveals very few differences in mathematics score gains and levels. Although the same holds true for reading two years prior to a student enrolling in SES, we find that students that enrolled but did not attend SES performed noticeably worse in the year prior to enrollment than did students that enrolled and attended SES. If this downward trajectory in test performance is associated with test measurement error then students that enrolled in but did not attend SES achievement are likely

to experience larger than expected gains during the following school year as their test scores return to historical average.

Second, these results may be explained, in part, by parents' opportunistic use of SES. Anecdotally, some parents enroll their child in SES to induce desired behavior in their child; that is, these parents threaten to send their child to after-school tutoring if their child doesn't improve their academic performance and/or change their attitude toward school. In a less draconian manner, parents may enroll their child with an SES provider as a precautionary measure (if their child's academic performance continues to slip their child can start attending after-school tutoring immediately) which can also motivate students to improve their academic performance.[11] Another plausible explanation is that not accounting for the content area in which a student receives tutoring biases estimates, which is the focus of the analysis reported in the next section.

6.2. *The Effect of SES on Student Test Score Gains by Content Area of Tutoring*

Even though the plurality of students (n=655) received tutoring in both reading and mathematics, 456 students received tutoring in reading only and 309 students received tutoring in mathematics only. The measured impact of SES on mathematics (or reading) test score gains will be biased downward when using the entire SES population because students who received tutoring in reading (or mathematics) will dilute the treatment effect. We fit a more precise specification of our

[11] Students that enrolled in but did not attend SES received out of school remedial services from another venue, although this explanation is less likely given the social and economic conditions of the district and the group of students under study. For example, 92.7 percent of students that enrolled with a SES provider but did not attend a single tutoring session qualified for free lunch program (7.3 percent quality for reduce-price lunch program). And, as reported by the United States Department of Agriculture (2008), children from families with incomes at or below 130 percent of the poverty level are eligible for free meals. Those with incomes between 130 percent and 185 percent of the poverty level are eligible for reduce-price meals, for which students can be charged no more than 40 cents. For the 2008-09 school year, 130 percent of the poverty level is $27,560 for a family of four (185 percent is $39,220) (USDA, 2009).

preferred modeling strategy by incorporating data on the content area of tutoring received by a student.

As displayed in Table 4, we introduced an additive term to distinguish the effects of receiving tutoring in only one subject, or both subjects compared to only one subject. Model (2), for example, measures the effect of receiving SES in mathematics only compared to receiving SES in reading only. If a student is tutored in mathematics only, he receives the sum of the betas, which is the equivalent to a 0.11 standard deviation increase in his mathematics score. The alternative scenario, which is reported by Model (3), indicates receiving tutoring in mathematics only results in a positive impact on test score gains in mathematics.

[Insert Table 4 Here]

Estimates displayed in Panel A of Table 4 further indicate that accounting for tutoring content area provides a more accurate measure for the effect of SES on student test score gains. Model (4), for example, compares the effects of SES in both subjects to the alternative of receiving tutoring in one subject only. Note that the size of the SES effect in mathematics is slightly larger than the baseline effect displayed in Model (1). The estimates in Panel B for reading display coefficients in the expected direction, but with too much variation to gain statistical significance at conventional levels.

There is error in the content area of tutoring measure. We do not know the relative weight placed on each subject for students receiving tutoring in both mathematics and reading, nor can we be absolutely certain that a student registered for tutoring in mathematics received tutoring only in that subject. While tutors may offer academic support in subjects beyond those defined in their students' individual learning plan, more precise measures of the SES treatment effect are obtained by limiting the regression sample to students who received tutoring in the subject of interest. Interestingly, we also find that students do not experience positive spillover effects in reading (or

mathematics) if they receive tutoring in mathematics (or reading) only, which offers additional support of an actual association between SES and student achievement.

6.3. *The Effect of SES on Student Test Score Gains by Student Attendance at SES*

We explore additional ways to measure the effect of SES on student test scores using data on student attendance as the indicator of interest. In select model specifications, student attendance is captured as a continuous variable measuring the total number of tutoring hours served. Other specifications used the percent of the allocated tutoring sessions attended which is also expressed as a continuous variable.[12]

As displayed in Table 5, there are moderate to large SES effects on student test score gains in mathematics and small to moderate effects in reading. A student that attends the mean number of tutoring hours is expected to increase his mathematics test score gain by 0.095 standard deviations, while a student at the 95th percentile of total hours attended is projected to gain almost one-fourth of a standard deviation more than expected. Although estimates reported in Model (3) indicate a weaker but significant relationship between the total hours of SES attended and student test score gains in reading, it suggests a student that attends the mean number of tutoring hours is expected to increase his reading test score gain by 0.072 standard deviations.

[Insert Table 5 Here]

Estimates measured by the percent of available hours attended are similar in that a greater intensity of tutoring is associated with greater standardized test score gains, but the magnitudes of the estimates are noticeably larger. For example, a student that attends the mean percentage of available tutoring is expected to experience an additional gain of 0.134 standard deviations in their

[12] We also ran regressions investigating the effect of a quadratic form of total hours attended to determine if the positive impact of attendance decreased with its growth, but the quadratic expression was statistically insignificant.

mathematics test score, which increases to 0.20 standard deviations for a student at the 95th percentile of the SES attendance rate distribution.

Estimates reported in Models (2) and (4) may be biased upward if the percentage of available SES hours attended proxies for unobserved time variant characteristics of an individual student or that individual student's school or home situation. For example, if student attributes such as student motivation or academic press of parents varies from one year to the next, the coefficient on the percentage of available SES hours attended is likely to capture at least part of the effect. Even so, the analysis finds that the total number of hours attended is significantly related to standardized test score gains in mathematics and reading, which further confirms the positive effect of SES.

6.4. *Moderators of the Effect of SES on Student Test Score Gains*

By introducing a simple interaction term between a binary variable for attended SES and a binary variable for a student characteristic, we conducted a series of exploratory analyses on the relationship between student characteristics and the estimated effect of SES on student test score gains. We first examine the heterogeneity of effects by student race, which is a salient topic given the large and persistent gap in test scores between students of different racial backgrounds (Clotfelter, Ladd, Vigdor, 2006; Stiefel, Schwartz, and Ellen, 2007; Hanushek and Rivkin, 2006), and then turn attention to LEP status, special education status, gender, and, finally, prior ability.[13]

As displayed in Table 6, SES does not appear to disproportionately benefit a particular racial group with the exception of the estimates from Model (6), which indicates that the test score gains of Hispanic students that received tutoring in mathematics increased, on average, 0.15 standard

[13] We do not investigate differential effects by socioeconomic status because a student must be eligible for free lunch or reduced-price lunch to enroll with a provider, and more than 90 percent of students enrolled in SES qualified for free lunch.

deviations more in mathematics than non-Hispanic students enrolled in SES. Further, there are notable changes in the direction of the sign and magnitude of the coefficient when accounting for the content area in which a student received services.

[Insert Table 6 Here]

We explored a similar specification for students identified as LEP. As displayed in Table 7, LEP students do not benefit more or less in mathematics from SES when compared to non-LEP students that attended SES. Although estimates reported in Models (15) and (16) demonstrate that they performed significantly worse in reading than non-LEP students, a negative effect is not unexpected. LEP students are still developing their English proficiency and, according to several SES coordinators, SES is a means to further develop language proficiency. Indeed, a simple logit model not only predicts LEP students are much more likely to enroll in reading-focused tutoring, but also that they are more likely to register and attend SES than their non-LEP peers.

We also explored a similar specification for students receiving special education services. As displayed in Models (5) and (6), students receiving special education services experience test score gains in mathematics approximately two-times larger than the expected gain for non-special education students enrolled in SES. There are even larger differences when test score gains in reading is the outcome variable. However, the large variability in the sampling distribution of special education students results in the differences being insignificant.

[Insert Table 7 Here]

Another set of analyses examined differential SES treatment effects by student gender. We find surprisingly strong effects on mathematics tutoring for female students. For example, as reported in Models (9) and (10) of Table 7, female students experienced an average test score gain increase in mathematics of approximate 0.12 standard deviations and .18 standard deviations,

respectively. There were negligible differences in test score gains among male and female students that attended SES as displayed in Models (11) and (12).

The last series of analyses examined whether tutoring services differentially benefitted lower- or higher-achieving students. Students were categorized into quartiles based on their prior year's performance in that subject. This allows for the measurement of the SES treatment effect on a specific group of students achieving at similar levels. As displayed in Table 8, we do not find a differential effect of SES by a student's prior level of achievement. Moreover, the estimates were not sensitive to the categorization of prior ability (i.e., terciles, quintiles, or deciles).

[Insert Table 8 Here]

In total, SES tutoring does not appear to disproportionately benefit a particular ethnic group or ability level. However, female students and students with disabilities disproportionately benefit from participating in SES. These findings are qualified by the fact some of the sample sizes creep toward the lower bound before stable parameter estimates become difficult to achieve. Thus it is important for readers to interpret student subgroup findings cautiously.

6.5. *The Cumulative Effects of SES Participation on Student Test Score Gains*

We also examined if there is a student-SES maturation effect. That is, does a student's exposure to SES from one year to the next affects that student's performance on the mathematics or reading assessments? We construct two binary variables denoting whether a student is enrolled in SES for the first time or is enrolled in SES for a second year. Perhaps strikingly, even after accounting for the matriculation of students and movement of schools in and out of eligibility, very few students attended SES for more than one year.

As displayed in Table 9, there is a large, cumulative effect of attending SES. Students attending SES for more than one year experience, on average, a cumulative increase of 0.39 standard

deviations and 0.48 standard deviations in their mathematics and reading test score gain, respectively. These estimates of the cumulative effects of SES align with those reported by Zimmer et al. (2007), as well as the evidence reported in evaluations of pre-NCLB after-school tutoring initiatives (Welsh et al., 2002) and summer school programs (Borman et. al., 2002).

[Insert Table 9 Here]

6.6. *Robustness Check*

To explore the robustness of our preferred estimation strategy, we first compare the performance of students enrolled in SES to future SES participants and then use propensity score analysis to estimate the SES treatment effect. In terms of the former strategy, we identify future participants defined as those students who were not yet eligible for SES but elected to enroll with a provider when they became eligible in the following school year. Both the movement of schools in and out of eligibility from year to year and the matriculation of students from elementary to middle schools allows for the formation of these future participants.

Table 10 displays estimates from the future participants modeling strategy. The estimates of the average SES treatment effect are slightly larger in magnitude when compared to estimates reported for the baseline model. Test score gains in mathematics are approximately 0.18 standard deviations greater for students enrolled in SES. Gains in reading are positive but still statistically indistinguishable from zero at conventional levels of significance. Assuming the timing of SES participation is not related to a student's test score in the year immediately preceding enrolling in SES, the future participants modeling strategy confirms estimates reported for our preferred modeling strategy.

[Insert Table 10 Here]

We also examine estimates of the SES treatment effect using propensity score analysis. As displayed in Panels C and D of Table 10, there is a significant effect of SES on test score gains in mathematics when measured by either the propensity score weights or stratification approach. Using the propensity scores as weights in a regression equation predicts slightly smaller estimates of the average SES treatment effect than the future participant analysis, while adjusting for covariates through stratification produces noticeably smaller estimates (e.g., estimates using weights are, on average, twice the size of the stratification estimates). Moreover, we find a weak negative effect of SES on reading test score gains, which is markedly different from the estimates generated by analytic strategies throughout this study even though readings scores tended to be statistically indistinguishable from zero in virtually all models.

Table 11 reports estimates from a series of robustness checks to the choice of gain specification. Models (3) to (8) and (13) to (18) enter prior achievement on the right-hand side of the regression equation, where select specifications model prior achievement as a linear, quadratic or cubic form. The dependent variable used in Models (9), (10), (19), and (20) was constructed by dividing the distribution of the students' prior year assessment scores into 20 equal intervals and then calculation the mean and standard deviation of the test score gain for all students starting in a particular interval. A student's test score gain was standardized by taking the difference between that student's nominal gain and the mean gain of all students in the interval over the standard deviation of all student gains in the interval.[14] For the most part, estimates models using alternative gain specifications are similar to those produced by our preferred modeling strategy.

[Insert Table 11 Here]

7. Conclusion

[14] This approach is described in Hanushek et al (2005) and has been used by Springer (2007, 2008) and others.

In this study, we examined the effect of SES on test score gains in mathematics and reading and whether particular subgroups of students benefit more from tutoring services. Our preferred analytic strategy employs a differences-in-differences estimation strategy, conditional on student, grade and time fixed effects. We explored the robustness of these estimates by comparing student test score gains between current and future SES participants. We also estimated the average SES treatment effect using propensity score analysis.

Approximately one in five students eligible to receive SES actually signed-up for tutoring services in a given school year. The average SES take-up rate falls from 21.0 percent of eligible students to 14.7 percent of eligible students after limiting the sample to students who attended at least one single tutoring session. Patterns of student attendance among those students enrolled in SES vary considerably (i.e., the average student attends 67 percent of his tutoring sessions), which lends support for examining the effect of SES on student test score gains by rate of attendance.

In terms of the average effect of SES on student test score gains, our preferred analytic strategy consistently detects significant and positive effects of SES attendance on test score gains in mathematics (0.09 standard deviations). Results in reading tended to be insignificant, however. These findings are maintained when controlling for the percent and absolute hours of SES tutoring sessions attended, with large positive effects for students receiving greater than 99 percent of their allocated tutoring in mathematics (0.33 standard deviations), but no significant impact on reading scores.

Our results further suggest that more precise measures of SES impact can be obtained by limiting the modeling sample to students who received tutoring in the subject of interest. We also find that the SES treatment effect also increases in the expected direction when controlling for the content area of tutoring. Specifically, students receiving tutoring in mathematics experience score

increases of 0.11 standard deviations, while those students receiving tutoring in reading experience score increases of 0.09 (but without statistical significance).

SES is measured to have a significant cumulative impact on students, with mathematics effect sizes of up to 0.39 standard deviations and 0.48 standard deviations for students receiving two or more years of tutoring in mathematics and reading, respectively.[15] SES also appears to disproportionately benefit female students and students with disabilities. At the same time, we do not find evidence of SES benefiting students belonging to a particular ethnic group or ability level, where ability level is determined by the previous year's quartile performance.

The measured, baseline treatment effect aligns with the findings from comparing students who reveal an SES enrollment preference through future participation with their pre-enrollment gains. This specification reveals statistically significant, positive effects of SES tutoring on student test score gains in mathematics (0.13 standard deviations), and small, positive effects that are not statistically significant in reading. In total, both approaches – which control for selection bias in different ways – find moderate, statistically significant effects on mathematics test score gains and positive, but statistically insignificant, results in reading test score gains.

Our findings are qualified on several dimensions. First, observational studies of this nature are always susceptible to selectivity bias or other spurious relationships arising from events or processes such as differential teacher effects. Second, we do not know the extent to which a student actually received academically focused tutoring when he attended SES, nor do we know the quality of those services. Third, even though most SES providers serve clients in urban school systems, it is unclear if students enrolled in tutoring services and the services offered in the district under study are similar to those of other districts throughout the nation.

[15] The upper and lower 95 percent confidence interval bounds are .56 and .21 in mathematics and .68 and .28 in reading.

8. References

Agondini, R. and Dynarski, M. (2004). Are experiments the only option? A look at dropout prevention programs. *Review of Economics and* Statistics, 86(1), 180-194.

Aronson, J., Zimmerman, J., and Carlos, L. (2002). Improving Student Achievement by Extending School: It is Just a Matter of Time. Washington: WestEd.

Blau, D. & Robins, P. (1988). Child-Care Costs and Family Labor Supply. *Review of Economics and Statistics, 70* (3), 374.

Borman, G. D., Rachuba, L. T., Fairchild, R., and Kaplan, J. (2002). *Randomized evaluation of a multi-year summer program: Teach Baltimore: Year 3 report.* Retrieved from http://www.childcareresearch.org/location/3333.

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

Catalano, R., Berglund, M.L., Ryan, J., Lonczak, H.S., and Hawkins, J.D. (1999). Positive Youth and Development in the United States: Research Findings on Evaluations of Positive Youth Program. Social Development Research Group. University of Washington.

Center on Education Policy. (2006). From the Capital to the Classroom: Year 4 of the No Child Left Behind Act. Washington, D.C.: Center on Education Policy.

Center on Education Policy. (2007). *State Implementation of Supplemental Educational Services under the No Child Left Behind Act*. Washington, D.C.

Chicago Public Schools: Office of Research, Evaluation, and Accountability (2005). *SES Tutoring Programs: An Evaluation of Year 3 in the Chicago Public Schools*. Chicago: Chicago Public Schools.

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

Clotfelter, C.T., Ladd, H.F., and Vigdor, J.L. (2007). The Academic Achievement Gap in Grades 3 through 8. *Review of Economics and Statistics.*

Connelly, R. (1992). The Effect of Child Care Costs on Married Women's Labor Force Participation. *Review of Economics and Statistics, 74* (1), 83.

Diaz, J.J. and Handa, S. (2006). An assessment of propensity score matching as a nonexperimental impact estimator. *Journal of Human Resources*, XLI(2), 319-345.

Dynarski, M., James-Burdumy, S., Moore, M., Rosenberg, L., Deke, J., and Mansfield, W. (2004). *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program: New Findings.* Washington, D.C.: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance.

Gelbach, J. (2002). Public Schooling for Young Children and Maternal Labor Supply. *American Economic Review, 92*(1), 321.

Gill, B. and Deke, J. (2008). Evaluating the Achievement Impact of Supplemental Educational Services Under No Child Left Behind. Presentation delivered at the 2008 Annual Meeting of the Association for Public Policy Analysis and Management.

Hanushek, E.A., and Rivkin, S.G. (2007). School quality and the black-white achievement gap. NBER Working Paper No. W12651.

Heinrich, C.J., Meyer, R.H., and Whitten, G. (2007). Supplemental education service under No Child Left Behind: Who signs up, and what do they gain? Wisconsin Center for Education Research Working Paper.

Heistad, D. (2005). *Analysis of 2005 supplemental education services in Minneapolis public schools: An application of matched sample statistical design.* Minneapolis, MN: Minneapolis Public Schools.

Hollister, R. (2003). The growth in after-school programs and their impact. Washington, DC: Children's Roundtable.

Kane, T. (2004). The impact of after-school programs: Interpreting the results of four recent evaluations. New York: William T. Grant Foundation.

Karweit, N. (1985). Should we lengthen the school term? *Educational Resesearcher*, 14(6), 9-15.

Lunceford, J.K. and Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine*, 23(19), 2937-2960.

Potter, A., Ross, S.M., Paek, J., McKay, D., Ashton, J., and Sanders, W.L. (2007). *Supplemental educational services in the State of Tennessee: 2005 - 2006*. Center for Research in Educational Policy. Memphis, TN.

Reisner, E.R., White, R.N., Brimingham, J., and Welsch, M. (2002). Building quality and supporting expansion of after-school projects: Evaluation rsults from the TASC after-school program. Washington, DC: Policy Studies Associates, Inc.

Rosenbaum, P.R. and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.

Springer, M.G., Pepper, M.J., Gardner, C., Bower, C.B. (2009). Supplemental Educational Services Under No Child Left Behind. In M. Berends, M.G. Springer, D. Ballou, H.J. Walberg (Eds., 2009). *Handbook of Research on School Choice*. New York: Routledge.

Shaul, M. S. (2006). *No Child Left Behind Act: Education Actions Needed to Improve Local Implementation and State Evaluation of Supplemental Educational Services. Report to Congressional Requesters.* Washington, DC: Government Accountability Office, Washington, DC.

Stiefel, L., Schwartz, A.E., and Ellen, I.G. (2007). Disentangling the Racial Test Score Gap: Probing the Evidence in a Large Urban School District. *Journal of Policy Analysis and Management*, 26(1), 7-30.

Wilde, E.T. and Hollister, R. (2002). How Close Is Close Enough? Testing Nonexperimental Estimates of Impact against Experimental Estimates of Impact with Education Test Score as Outcomes. Madison, Wisconsin: Institute for Research on Poverty. Downloaded from http://www.irp.wisc.edu/publications/dps/pdfs/dp124202.pdf.

Zimmer, R., Christina, R., Hamilton, L.S., and Prine, D.W. (2006). Evaluation of Two Out-of-School Program in Pittsburgh Public Schools: No Child Left Behind's Supplemental Educational Services and State of Pennsylvania's Educational Assistance Program. RAND Working Paper. Santa Monica, CA: RAND Corporation.

Zimmer, R., Gill, B., Razquin, P., Booker, K., Lockwood, J.R., Vernez, G., Birman, B.F., Garet, M.S., and O'Day, J. (2007). State and Local Implementation of the No Child Left Behind Act: Volume I - Title I School Choice, Supplemental Educational Services, and Student Achievement. *U.S. Department of Education.* Available at http://www.rand.org/pubs/reprints/2007/RAND_RP1265.pdf.

Figure 1. Percentage of Eligible Student Population Enrolled in SES by School and Year

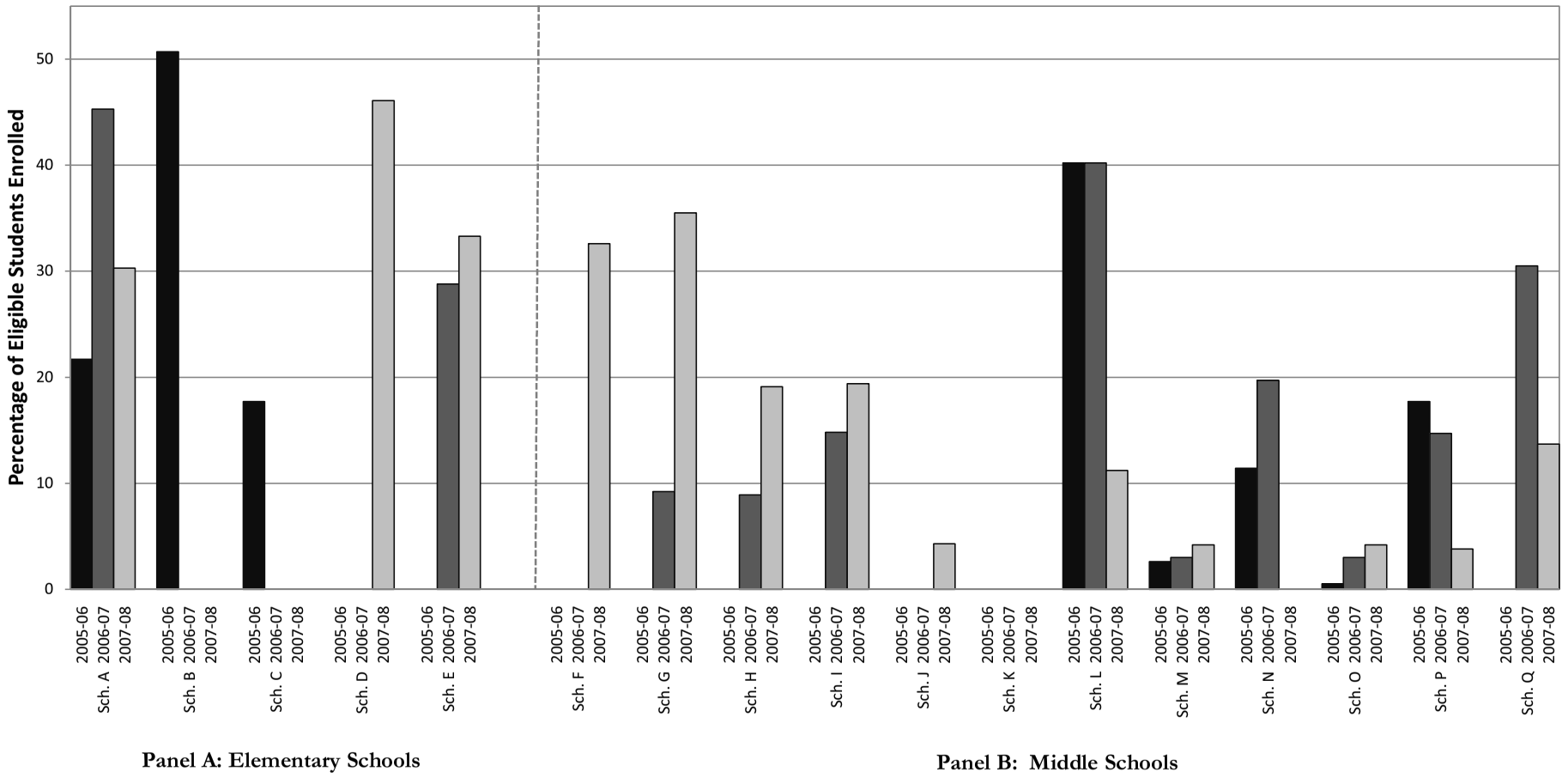

Table 1. Select Summary Statistics of Supplemental Educational Services

	School Year			All Years
	2005-06	2006-07	2007-08	(2003-04 to 2007-08)
District and Provider Information				
# of Students in District, Grades 3-8	28,484	28,862	29,075	143,801
# of Schools Required to Offer SES	6	12	14	17
# of SES Providers	8	13	14	20
Top 50% of Providers have x% of Market	85.1%	90.3%	89.8%	90.5%

	School Year			SES Years
	2005-06	2006-07	2007-08	(2005-06 to 2007-08)
Student Eligibility, Enrollment, and Attendance				
# of Eligible Students	1,432	3,904	4,422	9,758
% of Total Students who are Eligible	5.0%	13.5%	15.2%	11.3%
# of Students Signed Up	264	897	893	2,054
% of Eligible Students Signed Up	18.4%	23.0%	20.2%	21.0%
# of Students Receiving SES	194	583	657	1,434
% of Signed-Up Student Attended	73.5%	65.0%	73.6%	69.8%

Note: Data limited to continuously enrolled students in grades 3 through 8.

Table 2. Select Characteristics of Students in Grades 3 through 8, 2003-04 - 2007-08 School Year

	All Students		Schools Required to Offer SES						Students Enrolled in SES				SES Content Area					
			All Students		*Eligible Students*		*Enrolled Students*		*Attd. > 0 Hrs.*		*Did Not Attd.*		*Mathematics*		*Reading*		*Both*	
Gender	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Female	56,682	49.3%	5,935	48.8%	4,831	49.0%	1,088	51.2%	771	51.2%	317	51.1%	174	56.3%	232	51.0%	325	49.6%
Male	58,296	50.7%	6,225	51.2%	5,030	51.0%	1,038	48.8%	735	48.8%	303	48.9%	135	43.7%	224	49.0%	330	50.4%
Race/Ethnicity																		
Asian	3,730	3.2%	265	2.2%	192	1.9%	27	1.3%	17	1.1%	10	1.6%	<10	…	<10	…	<10	…
Black	54,939	47.8%	6,317	51.9%	5,486	55.6%	1,288	60.6%	906	60.2%	382	61.6%	186	60.2%	236	51.8%	415	63.4%
Hispanic	14,235	12.4%	2,218	18.2%	2,060	20.9%	437	20.6%	309	20.5%	128	20.6%	56	18.1%	153	33.6%	99	15.1%
Nat. Am / Pac. Is.	231	0.2%	16	0.1%	12	0.1%	<10	…	<10	…	<10	…	<10	…	<10	…	<10	…
White	41,843	36.4%	3,344	27.5%	2,111	21.4%	370	17.4%	272	18.1%	98	15.8%	59	19.1%	60	13.2%	137	20.9%
Other Characteristics																		
Reduced Price Lunch	9,327	8.1%	978	8.0%	978	9.9%	175	8.2%	132	8.8%	43	6.9%	36	11.7%	34	7.5%	59	9.0%
Free Lunch	61,544	53.5%	8,883	73.1%	8,883	90.1%	1,871	88.0%	1,321	87.7%	550	88.7%	261	84.5%	410	89.9%	572	87.3%
Limited English Proficient	6,623	5.8%	1,101	9.1%	1,044	10.6%	280	13.2%	190	12.6%	90	14.5%	28	9.1%	85	18.6%	75	11.5%
Special Education																		
< 5 Hours	3,399	3.0%	308	2.5%	248	2.5%	61	2.9%	43	2.9%	18	2.9%	<10	…	12	2.6%	19	2.9%
5 to 20 Hours	6,626	5.8%	779	6.4%	704	7.1%	200	9.4%	141	9.4%	59	9.5%	29	9.4%	28	6.1%	76	11.6%
> 20 Hours	3,355	2.9%	531	4.4%	463	4.7%	96	4.5%	63	4.2%	33	5.3%	10	3.2%	15	3.3%	38	5.8%
Average Hours	14		16		17		15		15		16		15		15		16	
Enrollment																		
Grade 3	20,544	17.9%	544	4.5%	502	5.1%	228	10.7%	199	13.2%	29	4.7%	17	5.5%	74	16.2%	69	10.5%
Grade 4	20,022	17.4%	550	4.5%	486	4.9%	180	8.5%	150	10.0%	30	4.8%	16	5.2%	44	9.6%	53	8.1%
Grade 5	18,934	16.5%	2,721	22.4%	2,311	22.8%	523	24.6%	378	25.1%	145	23.4%	89	28.8%	104	22.8%	183	27.9%
Grade 6	18,855	16.4%	2,849	23.4%	2,306	23.1%	455	21.4%	317	21.0%	138	22.3%	62	20.1%	94	20.6%	154	23.5%
Grade 7	18,544	16.1%	2,805	23.1%	2,211	22.2%	384	18.1%	239	15.9%	145	23.4%	71	23.0%	74	16.2%	94	14.4%
Grade 8	18,079	15.7%	2,691	22.1%	2,045	20.6%	356	16.7%	223	14.8%	133	21.5%	54	17.5%	66	14.5%	102	15.6%
Attendance Rate																		
School	95.7%		94.6%		94.4%		94.9%		95.1%		94.3%		94.7%		95.2%		95.1%	
SES	…		…		…		…		63.97%		…		60.23%		66.71%		64.17%	
Number of Students	114,978		12,160		9,861		2,126		1,506		620		309		456		655	

Notes: Authors own calculations. Authors suppress actual values in cells with less than 10 student observations (# = < 10) and (% = …).

Table 3. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Student Enrollment and Student Attendance

	Panel A: Mathematics			Panel B: Reading		
(model)	(1)	(2)	(3)	(4)	(5)	(6)
Registered	0.0766 (.0246) ***	0.0769 (.0246) ***	0.0880 (.0251) ***	0.0613 (.0258) **	0.0617 (.0258) **	0.0758 (.0265) ***
Sample Size	87355	87355	82280	85351	85351	80337
R^2	0.3463	0.3470	0.3667	0.3698	0.3699	0.3837
(model)	(7)	(8)	(9)	(10)	(11)	(12)
Attended	.0769 (.0294) **	.0696 (.0294) **	.0879 (.0299) ***	.0294 (.0306)	.0297 (.0306)	.0385 (.0313)
Sample Size	97323	87323	82248	85319	85319	80305
R^2	0.3464	.3470	.3666	.3698	.3700	.3836
Student Level Covariates		√	√		√	√
School Level Covariates			√			√
Student FE	√	√	√	√	√	√
Grade*Year FE	√	√	√	√	√	√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio.

Table 4. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Content Area of Tutoring

	Panel A: Mathematics				Panel B: Reading			
	Baseline estimates from Table 3, Model (9)	Received tutoring in mathematics only	Received tutoring in reading only	Received tutoring in both subjects	Baseline estimates from Table 3, Model (12)	Received tutoring in mathematics only	Received tutoring in reading only	Received tutoring in both subjects
(model)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attended	.0879 (.0299) ***	-.0129 (.0527)	.0960 (.0576) *	.0927 (.0576)	.0385 (.0313)	.0854 (.0558)	-.0028 (.0600)	.0897 (.0551)
Attended * Covariate (column)	… …	.1236 (.0785)	-.1001 (.0783)	.0546 (.0725)	… …	-.0816 (.0825)	.0969 (.0823)	-.0594 (.0708)
Sample Size	82248	81607 *	81607 *	81864	80305	79671	79671	80008
R^2	.3666	.3711	.3711	.3691	.3837	.3899	.3900	.3878
Student FE	√	√	√	√	√	√	√	√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio.

Table 5. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Student Attendance Patterns

	Panel A: Mathematics		Panel B: Reading	
(model)	(1)	(2)	(3)	(4)
Total Hours Attended	.0060	…	.0041	…
	(.0018)***	…	(.0017)**	…
Percent of Hours Attended	…	.0021	…	.0009
	…	(.0005)***	…	(.0005)*
Sample Size	82248	82248	80305	80305
R^2	.3667	.3667	.3837	.3836
Student FE	√	√	√	√
SES Content Area	√	√	√	√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Table 6. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Student Race/Ethnicity

	Baseline Estimates: Table 3, Models (9) and (12)	Full Sample	Black		Hispanic		White	
(model)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mathematics								
Attended	.0879 (.0299) ***	.1433 (.0332) **	.0478 (.0462)	.1447 (.0535) **	.0895 (.0337) **	.1183 (.0363) **	.1010 (.0330) **	.1669 (.0370) **
Attended * Covariate (column)	… …	… …	.0685 (.0602)	-.0023 (.0679)	-.0075 (.0717)	.1503 (.0888) *	-.0719 (.0766)	-.1193 (.0828)
Sample Size	82248	81896	82248	81896	82248	81896	82248	81896
R^2	.3666	.3692	.3666	.3692	.3666	.3693	.3666	.3693
(model)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Reading								
Attended	.0385 (.0313)	.0666 (.0340) *	.0350 (.0497)	.0780 (.0538)	.0435 (.0348)	.0587 (.0383)	.0336 (.0347)	.0679 (.0375) *
Attended * Covariate (column)	… …	… …	.0057 (.0636)	-.0188 (.0691)	-.0262 (.0788)	.0370 (.0827)	.0260 (.0794)	-.0072 (.0886)
Sample Size	80305	80040	80305	80040	80305	80040	80305	80040
R^2	.3836	.3877	.3836	.3877	.3836	.3877	.3836	.3877
SES Content Area		√		√		√		√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Table 7. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Student Characteristics

	Baseline Estimates: Table 3, Models (9) and (12)	Baseline Estimates: Table 6, Models (2) and (10)	Limited English Proficient		Special Education		Male Only		Female Only		Female	
(model)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mathematics												
Attended	.0879 (.0299) ***	.1433 (.0332) **	.0894 (.0315) ***	.1394 (.0347) ***	.0688 (.0322) **	.1257 (.0360) ***	.0592 (.0453)	.1064 (.0509) **	.1148 (.0391) ***	.1762 (.0431) ***	.0524 (.0427)	.1025 (.0481) **
Attended * Covariate	…	…	-.0149	.0433	.1345	.1159	…	…	…	…	.0686	.0774
(column)		…	(.0979)	(.1158)	(.0847)	(.0918)	…	…	…	…	(.0592)	(.0661)
Sample Size	82248	81896	82248	81896	82248	81896	41696	41525	40552	40371	82248	81896
R^2	.3666	.3692	.3666	.3692	.3666	.3693	.3643	.3663	.3749	.3785	.3666	.3693
(model)	(13)	(10)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Reading												
Attended	.0385 (.0313)	.0666 (.0340) *	.0602 (.0325) *	.0863 (.0355) **	.0131 (.0337)	.0491 (.0367)	.0137 (.0482)	.0567 (.0517)	.0586 (.0402)	.0727 (.0444)	.0085 (.0448)	.0498 (.0482)
Attended * Covariate		…	-.2783	-.2371	.1787	.1201	…	…	…	…	.0580	.0302
(column)		…	(.1158) **	(.1220) *	(.0886) **	(.0954)	…	…	…	…	(.0621)	(.0676)
Sample Size	80305	80040	80305	80040	80305	80040	40691	40581	39614	39459	80305	80337
R^2	.3836	.3877	.3837	.3878	.3837	.3878	.3910	.3952	.3774	.3814	.3837	.3836
SES Content Area		√		√		√		√		√		√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Table 8. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Student Ability

	Baseline Estimates:		Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Table 3, Models (9) and (12)	Table 6, Models (2) and (10)								
(model)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mathematics										
Attended	.0879 (.0299) ***	.1433 (.0332) **	.1011 (.0774)	.1229 (.0824)	.0095 (.0418)	.0592 (.0464)	-.0145 (.0487)	-.0158 (.0568)	-.2317 (.1253)*	-.0140 (.1530)
Sample Size	82248	81896	17069	17964	19610	19514	20789	20718	24780	24741
R^2	.3666	.3692	.4954	.4840	.8285	.8313	.8268	.8270	.4899	.4925
(model)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Reading										
Attended	.0385 (.0313)	.0666 (.0340)*	.0352 (.0994)	.0451 (.1059)	.0192 (.0421)	.0753 (.0467)	.0150 (.0444)	.0512 (.0503)	.0129 (.1083)	.0054 (.1207)
Sample Size	80305	80040	16246	16157	19284	19193	20701	20640	24074	24050
R^2	.3836	.3877	.5104	.5134	.8389	.8397	.8224	.8233	.4785	.4778
SES Content Area		√		√		√		√		√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Table 9. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Subject Focus and Years of Participation

	Baseline Estimates: Table 3, Models (9) and (12)	Baseline Estimates: Table 6, Models (2) and (10)	Enrolled in Tutoring 1 Year Only		Enrolled in Tutoring 2 or More Years	
(model)	(1)	(2)	(3)	(4)	(5)	(6)
Mathematics						
Attended	.0879 (.0299) ***	.1433 (.0332) **	0.0690 (.0309) **	.1235 (.0344) ***	.0690 (.0309) **	.1237 (.0344) ***
Attended * Covariate (column)	… …	… …	… …	… …	.2655 (.1153) **	.2657 (.1261) **
Sample Size	82248	81896	82049	81725	82248	81896
R^2	.3666	.3692	0.3674	.3698	.3666	.3693
(model)	(7)	(8)	(9)	(10)	(11)	(12)
Reading						
Attended	.0385 (.0313)	.0666 (.0340) *	0.0134 (.0323)	.0390 (.0351)	.0133 (.0324)	.0390 (.0351)
Attended * Covariate (column)	… …	… …	… …	… …	.3703 (.1233) ***	.4459 (.1403) ***
Sample Size	80305	80040	80114	79872	80305	80040
R^2	.3836	.3877	0.3846	.3889	.3838	.3879
SES Content Area		√		√		√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Table 10. Estimated Effects of Supplemental Education Services on Mathematics and Reading Test Score Gains by Student Enrollment and Attendance Patterns

	Baseline Estimates: Table 3, Models (9) and (12)	Baseline Estimates: Table 6, Models (2) and (10)	Panel B: Current versus Future Participants		Panel C: Propensity Score (weight)		Panel D: Propensity Score (stratification)	
(model)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mathematics								
Attended	.0879 (.0299) ***	.1433 (.0332) **	.1343 (.0721) *	.1789 (.0775) **	.0828 (.0165) ***	.0919 (.0175) ***	.0477 (.0232) **	.0494 (.0252) **
Sample Size	82248	81896	2374	2260	8477	8370	8466	8111
R^2	.3666	.3692	.3493	.3647	.0514	.0517	.0487	.0502
(model)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Reading								
Attended	.0385 (.0313)	.0666 (.0340) *	.0390 (.0856)	.1004 (.0995)	-.0330 (.0191) *	-.0303 (.0205)	-.0080 (.0265)	-.0041 (.0284)
Sample Size	80305	80040	2203	2076	8370	8104	8370	8113
R^2	.3836	.3877	.3902	.4248	.0191	.0185	.0166	.0157
SES Content Area		√		√		√		√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Table 11. Robustness Checks to Choice of Specification

	Baseline Estimates: Table 3, Models (9) and (12)	Baseline Estimates: Table 6, Models (2) and (10)	Prior mathematics or reading test score (t-1)		Interaction of prior mathematics and reading test score (t-1)		Polynomials of prior mathematics and reading test scores (t-1) and interaction of mathematics and reading test score (t-1)		Standardized gain score (bins)	
(model)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mathematics										
Attended	.0879 (.0299) ***	.1433 (.0332) **	.0265 (.0181)	.0514 (.0202) **	.0377 (.0188) **	.0648 (.0209) ***	.0364 (.0184) **	.0648 (.0205) ***	0.1427 (.0442) ***	0.2370 (0.0491) ***
Sample Size	82248	81896	82248	81896	80818	80485	80818	80485	82248	81896
R^2	.3666	.3692	.8973	.8979	.8935	.8941	.8977	.8983	.4234	.4259
(model)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Reading										
Attended	.0385 (.0313)	.0666 (.0340) *	.0180 (.0186)	.0265 (.0202)	.0243 (.0191)	.0367 (.0208) *	.0273 (.0188)	.0228 (.0256)	0.0813 (0.0451) *	0.1300 (0.0492) ***
Sample Size	80305	80040	80305	80040	80219	79954	80219	79954	80305	80040
R^2	.3836	.3877	.8919	.8925	.8855	.8862	.8898	.8905	.4222	.4248
SES Content Area		√		√		√		√		√

*, **, *** Estimates statistically significant from zero at the 10%, 5%, and 1% levels, respectively.

Notes: All models contain student and school level covariates and controls for grade-by-year effects. Student level covariate is student attendance at school. School level covariates include percent black, percent Hispanic, percent economically disadvantaged, percent scoring proficient in mathematics, percent scoring proficient in reading, and student teacher ratio. SES Content Area indicates SES sample is restricted to students receiving tutoring in mathematics or both mathematics and reading (or reading or both reading and mathematics) when mathematics (or reading) is the dependent variable.

Appendix A. Summary Statistics for Student Test Scores

	All Students		Schools Required to Offer SES						Students Enrolled in SES				Content Area					
			All Students		*Eligible Students*		*Enrolled Students*		*Attd. > 0 Hrs.*		*Did Not Attend*		*Mathematics*		*Reading*		*Both*	
	Mean	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD
Mathematics Scale Score																		
Grade 3	474.59	(31.07)	467.34	(26.27)	466.88	(26.47)	463.83	(23.83)	464.02	(23.26)	474.69	(31.12)	474.21	(20.54)	461.13	(24.84)	463.04	(24.13)
Grade 4	488.64	(34.50)	483.94	(34.37)	482.92	(33.94)	474.78	(32.62)	475.63	(31.81)	488.74	(34.50)	478.20	(29.50)	472.90	(33.55)	473.11	(32.20)
Grade 5	500.23	(38.69)	492.30	(35.13)	489.65	(34.17)	485.99	(31.77)	486.73	(31.18)	500.51	(38.78)	488.99	(38.00)	485.85	(27.02)	484.92	(30.68)
Grade 6	513.62	(44.59)	504.48	(40.49)	501.26	(39.85)	495.06	(37.55)	497.13	(36.94)	513.90	(44.66)	494.43	(37.07)	500.69	(41.28)	494.41	(34.17)
Grade 7	523.92	(48.27)	513.50	(45.54)	509.19	(44.58)	493.95	(42.84)	494.47	(42.27)	524.30	(48.23)	496.03	(40.57)	495.79	(40.45)	495.57	(42.70)
Grade 8	533.47	(50.44)	521.60	(45.85)	516.78	(45.48)	502.71	(44.19)	503.51	(43.66)	533.84	(50.41)	504.78	(40.01)	505.86	(46.30)	502.29	(42.76)
Reading Scale Score																		
Grade 3	486.72	(31.86)	476.19	(26.57)	475.44	(26.72)	472.64	(25.89)	472.24	(26.48)	486.86	(31.88)	479.53	(19.47)	469.59	(29.29)	471.06	(27.16)
Grade 4	493.24	(36.97)	481.30	(35.24)	480.85	(34.83)	473.73	(31.14)	474.72	(31.18)	493.39	(36.98)	479.60	(32.32)	469.88	(32.69)	471.15	(39.83)
Grade 5	504.05	(38.74)	492.92	(37.87)	489.82	(37.79)	484.43	(38.30)	485.23	(37.82)	504.44	(38.66)	493.79	(34.80)	479.88	(38.95)	482.47	(40.33)
Grade 6	518.37	(42.89)	506.85	(40.04)	502.90	(39.84)	496.01	(40.49)	497.00	(40.76)	518.74	(42.83)	498.67	(32.94)	504.13	(35.87)	492.22	(43.80)
Grade 7	522.39	(40.98)	510.26	(40.43)	506.09	(40.80)	494.41	(42.65)	495.07	(42.88)	522.75	(40.84)	500.05	(37.26)	488.96	(44.80)	497.99	(44.45)
Grade 8	536.18	(38.31)	528.05	(37.22)	523.72	(37.25)	517.07	(33.53)	518.20	(30.25)	536.40	(38.35)	514.90	(31.85)	526.37	(32.02)	515.02	(27.81)
Mathematics Gain Score																		
Grade 4	16.27	(23.54)	24.08	(22.44)	24.18	(22.30)	23.64	(22.97)	24.40	(22.48)	16.21	(23.54)	19.22	(18.82)	21.63	(20.22)	23.96	(25.35)
Grade 5	15.58	(25.26)	14.71	(28.12)	14.79	(28.25)	14.07	(28.03)	14.19	(29.03)	15.61	(25.17)	10.81	(26.97)	12.08	(30.11)	17.18	(28.32)
Grade 6	18.23	(27.11)	17.98	(29.88)	17.62	(30.65)	20.11	(34.46)	19.71	(32.89)	18.20	(27.00)	17.28	(34.49)	17.96	(35.36)	22.43	(30.75)
Grade 7	14.82	(29.45)	14.34	(31.19)	14.83	(31.82)	12.16	(34.15)	12.75	(35.93)	14.85	(29.36)	12.93	(27.96)	13.12	(34.63)	13.44	(37.08)
Grade 8	13.09	(30.17)	14.39	(34.60)	15.08	(36.55)	17.26	(40.09)	18.96	(42.17)	13.02	(29.99)	18.40	(44.46)	20.05	(40.13)	19.53	(38.65)
Reading Gain Score																		
Grade 4	8.96	(24.67)	9.80	(26.83)	10.37	(27.48)	10.90	(25.54)	10.77	(23.86)	8.95	(24.68)	4.76	(18.08)	10.18	(31.38)	5.94	(29.40)
Grade 5	15.30	(26.63)	15.73	(30.69)	15.81	(31.52)	15.70	(32.17)	15.58	(31.82)	15.29	(26.52)	13.23	(29.57)	15.86	(33.17)	16.98	(33.63)
Grade 6	19.43	(28.32)	17.97	(32.18)	17.53	(33.19)	17.15	(34.65)	14.82	(34.16)	19.51	(28.20)	17.07	(30.06)	23.03	(36.82)	10.14	(33.41)
Grade 7	9.31	(28.51)	8.94	(32.21)	9.94	(33.75)	10.36	(40.35)	8.77	(39.02)	9.32	(28.35)	9.26	(33.45)	8.90	(32.92)	7.97	(43.48)
Grade 8	16.13	(27.30)	20.48	(31.68)	21.20	(34.16)	24.15	(35.32)	25.35	(34.45)	16.02	(27.18)	26.89	(40.99)	27.64	(34.52)	24.52	(31.98)

Faculty and Research Affiliates

Matthew G. Springer
Director
National Center on Performance Incentives

Assistant Professor of Public Policy and Education
Vanderbilt University's Peabody College

Dale Ballou
Associate Professor of Public Policy and Education
Vanderbilt University's Peabody College

Leonard Bradley
Lecturer in Education
Vanderbilt University's Peabody College

Timothy C. Caboni
Associate Dean for Professional Education and External Relations
Associate Professor of the Practice in Public Policy and Higher Education
Vanderbilt University's Peabody College

Mark Ehlert
Research Assistant Professor
University of Missouri – Columbia

Bonnie Ghosh-Dastidar
Statistician
The RAND Corporation

Timothy J. Gronberg
Professor of Economics
Texas A&M University

James W. Guthrie
Senior Fellow
George W. Bush Institute

Professor
Southern Methodist University

Laura Hamilton
Senior Behavioral Scientist
RAND Corporation

Janet S. Hansen
Vice President and Director of Education Studies
Committee for Economic Development

Chris Hulleman
Assistant Professor
James Madison University

Brian A. Jacob
Walter H. Annenberg Professor of Education Policy
Gerald R. Ford School of Public Policy
University of Michigan

Dennis W. Jansen
Professor of Economics
Texas A&M University

Cory Koedel
Assistant Professor of Economics
University of Missouri-Columbia

Vi-Nhuan Le
Behavioral Scientist
RAND Corporation

Jessica L. Lewis
Research Associate
National Center on Performance Incentives

J.R. Lockwood
Senior Statistician
RAND Corporation

Daniel F. McCaffrey
Senior Statistician
PNC Chair in Policy Analysis
RAND Corporation

Patrick J. McEwan
Associate Professor of Economics
Whitehead Associate Professor of Critical Thought
Wellesley College

Shawn Ni
Professor of Economics and Adjunct Professor of Statistics
University of Missouri-Columbia

Michael J. Podgursky
Professor of Economics
University of Missouri-Columbia

Brian M. Stecher
Senior Social Scientist
RAND Corporation

Lori L. Taylor
Associate Professor
Texas A&M University

NATIONAL CENTER ON

Performance Incentives

EXAMINING PERFORMANCE INCENTIVES IN EDUCATION

National Center on Performance Incentives
Vanderbilt University Peabody College

Peabody #43
230 Appleton Place
Nashville, TN 37203

(615) 322-5538
www.performanceincentives.org

VANDERBILT
PEABODY COLLEGE