Exceeds Expectations Incentive: Teacher Characteristics and Student Achievement 2011 Report Brief

Amy N. Farley
University of Colorado at Boulder
School of Education

Internal Report; DO NOT CITE WITHOUT AUTHOR'S PERMISSION.

SECTION 1: INTRODUCTION

This report brief assesses the degree to which the *Exceeding CSAP Expectations* (also referred to as simply Exceeds Expectations, or EE) incentive rewards teachers who best promote their students' learning. Specifically, I examine the differences between teachers who earn the incentive and those who are eligible but do not qualify. I also analyze achievement trends in mathematics and reading for students taught by EE teachers compared to students in other classrooms, conditioned on prior achievement. I first outline – in this section – the motivating policy concerns, research questions, and relevant background literature. Section 2 describes the data sources and measures used for the analysis, and Section 3 outlines the three analytic approaches utilized. Section 4 presents findings from each of the analytic approaches. Finally, in Section 5, I discuss conclusions, potential policy recommendations, and directions for future research.

In the last decade, teacher pay-for-performance reform efforts have gained traction both in the political community and within education reform initiatives. New programs are being implemented across the U.S. each year, particularly in large, urban districts; a ballot initiative requiring merit pay appeared on the 2008 Oregon ballot, although it ultimately failed to win enough votes; and in 2007 over $97,000,000 in federal funds were distributed for the development of pay-for-performance programs under the Teacher Incentive Fund. Despite broad support for pay-for-performance, once common element continues to be controversial: paying teachers for outstanding student achievement as measured by performance on state standardized tests.

Literature on pay-for-performance reform suggest two possible reasons that explicitly rewarding teachers for student achievement is so contentious. Some scholars contend that the main reason is politically motivated: “The main impediment to merit pay in public schools is the opposition of teacher unions,” who have traditionally rejected all pay-for-performance systems. Others argue that bonuses linked to student achievement are controversial for more technical reasons, including “the inherent complexity of designing systems that are valid and reliable [and can] consistently identify and reward the most effective teachers” (Buddin et al., 2007, p. 1). Although researchers generally find that teachers do affect the achievement of their students (Rivkin, Hanushek, & Kain, 2005; McCaffrey et al., 2003; Wright, Horn, & Sanders, 1997), few can agree on a consistent and methodologically defensible way to fully disentangle teacher-effects from student- and school-level factors (McCaffrey, Han, & Lockwood, 2008). Performance measures such as scores on state assessments are also vulnerable to multiple methodological limitations, including sampling error, measurement error, missing data, and inconsistencies in the interpretation of scaled scores across the distribution (Buddin et al., 2007; Patz, 2007).

Many of these problems are being addressed in research on pay-for-performance and current reform efforts. Educational researchers have begun to tackle methodological concerns in a rich body of literature. Furthermore, districts and policy-makers have worked with education researchers to broaden and refine measures of student achievement. Finally, sophisticated technical analyses – including value-added models of student achievement and other growth models, including the Colorado Growth Model – allow us to more confidently isolate teacher effects from other confounding variables.

One of the most compelling things about Denver's pay-for-performance plan is that supporters were able to garner teacher and union support, in addition to an endorsement from the greater Denver community demonstrated by the passage of a mill levy providing $25 million to the program. Goldhaber (2008) attributes this extensive support "to the willingness [of policy-makers] to engage the union and teachers from the beginning in thinking through the design and implementation of the proposed system" (p. 10). Some researchers have suggested that ProComp's architects were able to garner local union support because the reform includes incentives for several criteria other than increasing student achievement (Goldhaber, 2008; Gonring, Teske, & Jupp, 2007). In fact, the evidence strongly suggests plans that offer financial incentives for teachers working in hard-to-serve schools or hard-to-staff positions are more readily accepted by both unions and individual teachers than those that reward teachers for student achievement (Jacob & Springer, 2008; Goldhaber, DeArmond, & DeBurgomaster, 2007).

Despite its framing as a more holistic and comprehensive program, ProComp does provide an incentive for teachers whose students' achievement growth exceeds expected norms on the state standardized tests. Implemented in 2006-2007, the Exceeds Expectations component is the only element of ProComp that rewards *individual* teachers on the basis of standardized test scores; to earn the Exceeds Expectations incentive under ProComp, teachers must have 50 percent of eligible students within a single subject (reading, writing, or mathematics) at or above the 55th percentile for statewide student growth using Colorado's Student Growth Indicator.[1]

Consistent with existing research on pay-for-performance, the Exceeds Expectations incentive of ProComp is one of the most contentious components of the reform. According to 2007, 2008, 2009, and 2010 district-wide teacher and principal surveys, teachers and staff appear somewhat

---

1 There are additional eligibility criteria, for both teachers and students, that are outlined in Section 2 of this report.

skeptical about the efficacy and appropriateness of the reform, as evidenced both by their responses to forced choice survey questions and open-ended reflections on the incentive. Specifically, teachers have voiced concerns that the incentive is disproportionately earned by teachers in high-achieving schools and/or those who serve less disadvantaged student populations.

The questions that DPS teachers have about this incentive are significant. Not only do they jeopardize the teachers' continued support of the reform, but it also could negatively impact the potential for the incentive to be successful: If teachers feel that the incentive structure disproportionately awards teachers serving more affluent, privileged populations, they may be unlikely to change their behaviors to receive the monetary incentive. This apprehension compromises teacher buy-in: If teachers feel the incentive is inappropriate/unfair, they may be unlikely to change their instructional behaviors in order to receive the monetary incentive. [2] In moving forward with this high-profile policy issue, it is imperative that we explore the legitimacy of teachers' equity concerns, specifically regarding the potential over-identification of teachers serving non-disadvantaged students. In this report brief, I examine the EE incentive to explore teachers' concerns and evaluate the extent to which EE teachers do positively increase student achievement.

SECTION 2: DATA SOURCES & MEASURES

Data were collected for teachers eligible for the Exceeds Expectations incentive in 2007-2008 and 2008-2009.[3] Appropriately defining the population of teachers included in the analyses was critical: The sample must only include teachers who could have *actually* earned the EE incentive as a result of high student growth. To be eligible for the incentive, teachers must participate in ProComp, be a part of the DCTA (Denver Classroom Teachers Association), and be the primary teacher for their class; additionally, they must teach a qualified math or language arts course in grades 4 through 10, and must teach at least 10 qualified students in their subject area (students that have scores from the prior year, and who are enrolled for 85% of the course and attend 85% of the time) (Denver Public Schools Web Site, n.d.). In determining the sample for analyses, teachers were *excluded* who failed to meet any of the above eligibility criteria.

---

2 Prior research on ProComp has suggested just this: Internal report briefs from 2009 and 2010 suggest that teacher attitudes in Denver are strongly associated with reported behavior changes (Spindler, Subert, & Wiley, 2009; Spindler & Farley, 2010).

3 Data from 2009-2010 was unavailable during data collection, although may be reanalyzed in future work.

For 2007-2008, the final sample included 399 eligible teachers, and roughly 54 percent (N=217) earned the EE incentive in either mathematics or reading. For mathematics, 263 teachers were eligible for the incentive, and almost 48 percent earned it (N=126); in reading, 143 of the eligible 275 teachers (or 52 percent) received the incentive. In 2008-2009, the population of eligible teachers grew – which is consistent with the policy of compulsory ProComp participation for new hires – to 615 teachers. Forty-eight percent (N=298) of them earned the incentive in reading or mathematics. Of the 426 eligible teachers in mathematics, 46 percent earned the incentive (N=197); slightly fewer eligible teachers (40 percent, N=164 out of 409 eligible) earned the incentive in reading.

In research on schools, and particularly research examining student achievement, scholars agree that there are generally three levels of characteristics that affect student achievement and school effectiveness: student-, teacher-, and school-level factors. As such, student-, teacher-, and school-level data were collected for EE teachers and their colleagues who were eligible for the incentive but did not receive it. For the eligible sample of teachers – and their respective students – data were collected from two primary data sources: (1) student-level longitudinal achievement data in reading and mathematics; and (2) student-, teacher- and school-level demographic and descriptive data culled from district databases. In building the data, I drew on human resources teacher-level data, which I linked to (a) school-level data regarding teacher and school characteristics and (b) individual student-level achievement and demographic data. For teacher-level analyses, student-level data was aggregated to the classroom level to augment school-level demographic data.

In what follows, I outline briefly the school-, teacher-, and student-level measures collected and included in the analyses.

School-Level Variables

School-level variables included: (a) Hard-to-Serve school status, another incentive under the ProComp system, and an overall proxy for school-wide poverty; (b) Percentage of students schoolwide identified as minority, which include students identified as Black, American Indian, Hispanic, or Asian; [4] (c) Percentage of students schoolwide receiving ELL services; (d) Percentage of students schoolwide receiving free or reduced-priced (FRL) lunch; (e) Percentage

[4] I considered excluding Asian students from this category – as they often outperform other students of color on measures of student achievement – but trends remained consistent across schools and teachers in both cases.

of students schoolwide identified as gifted (GT); and (f) School level, including dummy variables for elementary, middle, and high schools.

Teacher-Level Variables

Teachers' effects on student achievement are often contested in the literature on student performance, with some scholars arguing that teacher can do little to close the achievement gap (Rothstein, 2004; Coleman, 1966) and others arguing that high-quality teachers can drastically affect student achievement (Rivkin, Hanushek, & Kain, 2005). Despite these claims, there is little consensus on what variables can actually predict high-quality teachers. As such, in addition to information about Exceeds Expectations status, I included only those variables that have repeatedly been linked to measures of teacher quality and increased student achievement in the literature. The variables included for this analysis include: (a) Overall Exceeds Expectations status, a dummy that simply characterizes teachers who earned the incentive in any subject; (b) A variable specifically designating those teachers who exceeded expectations in mathematics and/or in reading; (c) The percentage of students within reading and mathematics classes with a student growth percentile of 55 or higher, as required in the EE eligibility criteria;[5] (d) Total years of service in Denver Public Service; (e) Teacher status in the ProComp system, including either voluntary or compulsory participant (newly employed teachers must opt-in to ProComp); and (f) Masters degree or higher.

Student-Level Variables

Because teachers are the unit of analysis for the two of my analytic approaches, I am unable to use traditional student-level data. As such, for those analyses where teachers are the unit of analysis, student level data has been aggregated to the teacher level. For teachers with only one cohort or class of students (e.g. elementary school teachers), these values are aggregated at the individual class level. For teachers with multiple eligible classes of students, this data is aggregated across classes; this is matched to the aggregation process used in determining Exceeds Expectations status. These variables included: (a) Mean CSAP scores in reading and mathematics, both as scaled scores and also converted to grade-level z-scores to account for raw score differences across grades; (b) Percentage of students who qualify for FRL; (c) Percentage of students identified as minority; (d) Percentage of students who were "unsatisfactory," "proficient," and "advanced" on previous year CSAP reading and math tests; and, (e) Total number of students in a given teacher's class or classes. For student achievement analyses,

---

[5] Aggregate student growth percentile data was drawn from DPS data and originally generated by the Colorado Department of Education, using the Colorado Growth Model. The Colorado Growth Model is a normative measure of student growth that provides an estimate of student performance relative to students with similar score histories.

additional student-level data included individual (a) FRL status and (b) scaled scores on reading and mathematics CSAP tests.

SECTION 3: ANALYTIC APPROACHES

To address the primary research questions of interest, I utilized three primary analytic approaches, including (a) descriptive analyses of the differences and similarities between EE and non-EE teachers, (b) a probabilistic model of the likelihood of receiving the EE incentive for various teacher types, and (c) student-level analyses of student achievement patterns.

I first examine characteristics – at the school, teacher, and classroom/student level – of teachers identified under the reform, and attempt to quantify how Exceeds Expectations teacher differ from the general population of teachers. In doing this, I evaluate the effectiveness of the Exceeds Expectations incentive in isolating teacher effects from other student- and school-level characteristics. Buddin et al. (2007) claim that in order for pay-for-performance to be successful, performance measures must "account for differences in the student populations among classrooms and provide precise estimates that distinguish among teachers" (Buddin et al., 2007, p. 18). This is exactly the question I take up in the first two parts of this analysis.

I then employ binary logistic regression to predict the probability of receiving the EE incentive on the basis of a battery of school- and teacher- variables. These analyses provide a means for determining if, and to what extent, classroom- and student-level factors can predict a teacher's probability of receiving the reward.

Finally, achievement analyses explore the relationship between EE teachers, student achievement, and poverty. By relying on hierarchical modeling techniques, we can examine the nested nature of schools – with students nested in teachers, nested within schools. Results from these analyses allow us to consider the impact of EE teachers on student learning, and address the claim that the EE incentive only rewards teachers who serve students that are traditionally more successful in school.

SECTION 4: RESULTS

Descriptive Analyses – Teacher & Student Characteristics

Tables A-1 and A-2 (in Appendix A) present comparisons of EE teachers and nonEE teachers for the collected student-, teacher-, and school level variables for the 2007-2008 and the 2008-2009 incentive. Generally – across school types, CSAP subjects, and school years – these analyses suggest that EE teachers tend to teach fewer students of color and low-income students, and are more likely to work in schools with lower percentages of those populations. They also tend to teach slightly more students identified as gifted and talented and slightly higher proportions of students who previously scored at the proficient or advanced level; correspondingly, their classrooms contain slighter lower proportions of students who previously scored at the unsatisfactory level.

In general, the data suggests that patterns are consistent across elementary and secondary school teachers and in both 2007-2008 and 2008-2009; however, in comparing student populations, secondary EE teachers seem to have student demographics that more closely resemble their nonEE peers than within the elementary teacher population. Within secondary teachers, however, difference in prior achievement seem to be greater. This distinction – a greater disparity in demographic composition at the elementary level and in prior achievement at the secondary level – may be in part attributable to the instructional organization of schools at each level. While elementary schools have greater variation in student demographics across schools, students tend to be grouped within schools heterogeneously. On the other hand, secondary schools may more closely resemble one another in terms of schoolwide demographics, but within classroom variation – especially in terms of prior achievement – is increased due to common instructional and tracking practices.

Finally, demographic differences in student populations seem to be greater (when comparing EE and nonEE teachers than) for reading teachers than for math teachers; this may suggest that the impact of demographic characteristics is more easily overcome in mathematics than in reading achievement – a hypothesis that is supported in the academic literature.

Logistic Model – Probability of Receiving the EE Incentive

After descriptively examining differences between teachers who earned the EE incentive and those who did not, I model the receipt of EE on a battery of independent variables including teacher characteristics, school-level characteristics, and student-level characteristics aggregated to the classroom level.

Essentially, the logistic model of EE incentive receipt is specified as:

$$P(Y_t = 1) = \frac{e^{f(\underline{X})}}{1 + e^{f(\underline{X})}}$$

In the model above, Y_t is a dichotomous variable that takes a value of 1 if an eligible teacher earns the incentive, and 0 otherwise, and $\underline{X}$ represents a combination of variables, including years of service, advanced degree, working in a hard to serve school, the total number of students within a teacher's class(es), the percentage of teachers retained in a given teacher's school in the prior school year, and a dummy indicator for working in at the elementary level. The results from this analysis are presented in Appendix B.

A test of the full model with all predictors against a constant-only model was statistically reliable ($\chi 2 = 36.7$, $p < .001$). The full model also added predictive power, with 66.3% of Exceeds Expectation teachers accurately predicted, versus 54.8% in the constant-only model. Estimates from the specified model are provided in Table B-1 and help approximate the degree to which teacher and school characteristics are related to receipt of the EE award. In logistic regression, odds ratios can be computed by raising e to the power of the logistic coefficient; odds ratios are more easily interpretable than the logistic coefficient because they quantify the extent to which the odds of passage increase or decrease in relation to a given independent variable, holding other modeled variables constant. For example, the odds ratio for the dichotomous hard-to-serve variable is estimated at 0.5, meaning that the *odds* of earning the incentive for a teacher working in hard-to-serve school would be 0.5 times that of a teacher in a non-HTS school, holding all other variables constant.

These results reinforce the findings in the descriptive analyses outlined above, suggesting that working in a hard-to-serve school is a reliable predictor for EE designation. These findings must be interpreted cautiously, however: While they may suggest that teachers of low-poverty populations are disproportionately awarded the Exceeds Expectations incentive, it is unclear whether this pattern occurs because some teachers are unfairly disadvantaged or whether teachers of more advantaged students also happen to be more effective.[6]

Figure B-1 in Appendix B presents a graphical representation of the relationship between teacher years of service, employment in a hard-to-serve school, and the predicted probability of earning the EE incentive, based on the estimates of the model in the table above.[7] The figure presents two ogives – each mapping the predicted probability of EE award based on years of service. One line represents teachers working in schools designated hard-to-serve (HTS) and

---

6 This is a critical point; the differences between the student populations served by EE teachers and teachers eligible for the award

7 Values for Masters +, Number of Students, % Teacher Retained, and Elementary School Teacher were fixed at Masters = 0 (or no Masters), 40 students, 85 percent of teachers retained (the mean across the population), and Elementary = 1 (to designate an elementary teacher) to enable comparison of teachers in hard-to-serve school schools with those in non-hard-to-serve schools.

one for those in schools not designated hard-to-serve (non-HTS). This presents a dramatic representation of the effect of the percentage of students eligible for free or reduced-price lunch on predicted probability of Exceeds Expectations, with teachers in hard-to-serve schools less likely to receive the Exceeds Expectations incentive, holding other variables constant. Again, this finding must be interpreted cautiously, because it do not necessarily suggest that being in a hard-to-serve school unfairly disadvantages teachers from receiving the incentive.

Student-Level Achievement Analyses

Finally, additional analyses explored the nested nature of student achievement data through an hierarchical linear modeling approach. This allows for a more thorough examination of score gains for students in EE classrooms, while also accounting for student-, teacher-, and school-level demographics.

Although conceptually similar to OLS regression, the three-level model accounts for the nested structure of the data (students within teachers' within schools). The first level of the model allows us to evaluate student score trends on CSAP subjects (math or reading) based on prior scores and student FRL status. The second level of the model accounts for variation in teachers, and includes a variable to estimate the impact of receiving the EE incentive. The final level accounts for school-level variation and includes a hard-to-serve indicator.

The full three-level model is presented below:

Level-1 Model (Students): $\mathrm{CSAP}_{ijk} = \pi_{0jk} + \pi_{1jk}*(CSAP\_PRIOR_{ijk}) + \pi_{2jk}*(FRL_{ijk}) + e_{ijk}$

Level-2 Model (Teachers):

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}*(EE_{jk}) + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k} + \beta_{11k}*(EE_{jk})$$
$$\pi_{2jk} = \beta_{20k} + \beta_{21k}*(EE_{jk})$$

Level-3 Model (Schools):

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(HTS_k) + u_{00k}$$
$$\beta_{01k} = \gamma_{010} + \gamma_{011}(HTS_k)$$
$$\beta_{10k} = \gamma_{100} + \gamma_{101}(HTS_k)$$
$$\beta_{11k} = \gamma_{110} + \gamma_{111}(HTS_k)$$
$$\beta_{20k} = \gamma_{200} + \gamma_{201}(HTS_k)$$
$$\beta_{21k} = \gamma_{210} + \gamma_{211}(HTS_k)$$

Parameter estimates and model plots for reading and mathematics in grades 5, 8 and 10 are presented in Appendices D, E, and F. These results suggest that students within EE classrooms perform better on CSAP reading and math, conditioned on prior achievement. Although the score differences are small, they are statistically significant.

Furthermore, by examining trends for achievement conditioned on FRL status and schoolwide poverty, these analyses suggest that EE teachers are *not* less effective with disadvantaged students. In fact, in every case – across grades and subject – FRL students in EE classrooms score higher – as a function of prior achievement – than students in non-EE classrooms. Furthermore, the gap between FRL students and non-FRL students seems to be decreased in EE classrooms, and the model predicts higher scores for FRL students in EE classrooms than for non-FRL students in non-EE classrooms. Students within hard-to-serve schools also score higher in EE classrooms, conditioned on prior achievement, than their peers in non-EE classrooms in non-hard-to-serve schools.

SECTION 5: DISCUSSION, POLICY CONSIDERATIONS & NEXT STEPS

In general, the results presented in this report brief are mixed. On the one hand, EE teachers do differ in some important ways from other teachers who are eligible for the incentive. While this does not necessarily suggest that methodology is biased, it does raise questions about potential clustering of high quality teachers in less disadvantaged schools. It is worth pursuing additional analyses in this vein and considering the implementation of an incentive structure that more comprehensively addresses potential peer composition effects in student achievement. While the Colorado Growth model does account for score trajectories – essentially comparing students only to other students with similar achievement histories – it does not account for the inherent difficulty in teaching an entire class of disadvantaged or low-achieving students.

Despite these differences in teacher characteristics, results from the achievement analyses are more promising: EE teachers are more effective in raising student achievement across the board – a result that provides strong counter-evidence to the prevailing belief among district teachers that the incentive does not really reward teachers who are effective with disadvantaged populations. That students who are eligible for free- or reduced-price lunch and students in hard-to serve schools perform better than more advantaged peers in non-EE classrooms is a very compelling finding.

Future research should utilize additional years of data, as the sample size of teachers in both 2007-2008 and 2008-2009 is relatively small (N=399 and 615 respectively). Furthermore, analyses that utilize individual student growth percentiles may shed additional light on achievement patterns and trends for EE teachers. As the district continues to evaluate ProComp and the individual incentives, it may be valuable to consider two things. First, would an increased consideration of peer and composition effects allow for more teachers of disadvantaged students to receive recognition for raising student achievement? And, second, is the number and proportion of receiving EE incentive – nearly 50 percent of those who are truly eligible – consistent with the theory of action and intent of the incentive, or would it be prudent to consider rewarding fewer teachers with more exceptional student achievement gains? As was stated previously, the achievement difference in EE classrooms are significant, but they are also practically small. At any rate, the results of these analyses shed important light on a promising reform effort; taken together, they provide evidence that teachers who earn the EE incentive do effectively promote student learning. FRL students in EE classrooms score higher – as a function of prior achievement – than students in non-EE classroom

Appendix A.

Descriptive Statistics: Characteristics of Exceeds Expectations Teachers, Students, and Schools

Table A-1.

2007-2008 descriptive statistics by CSAP subject, school level, and exceeds expectations incentive receipt

	Math CSAP				Reading CSAP			
	Elementary		Secondary		Elementary		Secondary	
	NonEE	EE	NonEE	EE	NonEE	EE	NonEE	EE
Teacher Years of Service	8.9	11.5	11.3	10.6	7.8	12.2	8.4	11.3
% of Teachers in Hard to Serve Schools [8]	51.5	43.2	43.3	27.3	55.2	33.3	45.2	19.3
% ELL Students in School	24.2	17.2	9.6	12.1	17.3	13.8	15.5	13.6
% GT Students in School	17.0	24.3	20.4	22.1	15.0	27.8	16.6	23.6
% FRL Students in School	73.6	62.7	66.0	66.1	73.1	61.1	72.5	65.8
% Minority Students in School	84.7	73.0	77.6	78.1	83.0	72.5	83.4	77.6
% Teachers Retained in School	86.6	86.0	78.8	84.8	85.1	87.5	81.4	85.7
% FRL Students in Class(es)	76.0	63.2	67.0	62.5	73.9	61.6	71.1	59.0
% Minority Students in Class(es)	90.1	73.4	82.1	78.0	87.0	73.3	85.9	75.8
% "Proficient" in Prior Year - Class	31.1	32.1	19.1	22.8	37.6	41.5	30.0	37.5
% "Advanced" in Prior Year - Class	10.4	17.5	7.9	11.1	1.5	5.9	2.4	4.5
% "Unsatisfactory" in Prior Year - Class	11.7	11.5	24.2	24.2	14.9	14.5	21.2	16.9
% Students Met Growth Criteria	30.6	66.8	37.9	62.1	35.1	62.9	38.2	61.5
Mean CSAP Scale Score	463.5	506.2	520.8	538.2	566.3	602.2	600.8	634.8
Mean CSAP Prior Scale Score	440.5	459.3	516.6	515.5	537.5	564.8	581.0	605.1
Mean CSAP Z-Score	-0.285	0.219	-0.208	0.032	-0.180	0.232	-0.343	0.067
Mean prior CSAP Z-Score	-0.156	0.043	-0.102	-0.050	-0.192	0.092	-0.336	-0.062

8 Note that these figures can be misinterpreted. They represent the percentage of teachers who work in hard to serve (HTS) schools for both the populations that received the EE incentive and those who did not. It does NOT represent the percentage of people awarded the incentive within HTS schools versus non-HTS schools.

Table A-2.
2008-2009 descriptive statistics by CSAP subject, school level, and exceeds expectations incentive receipt

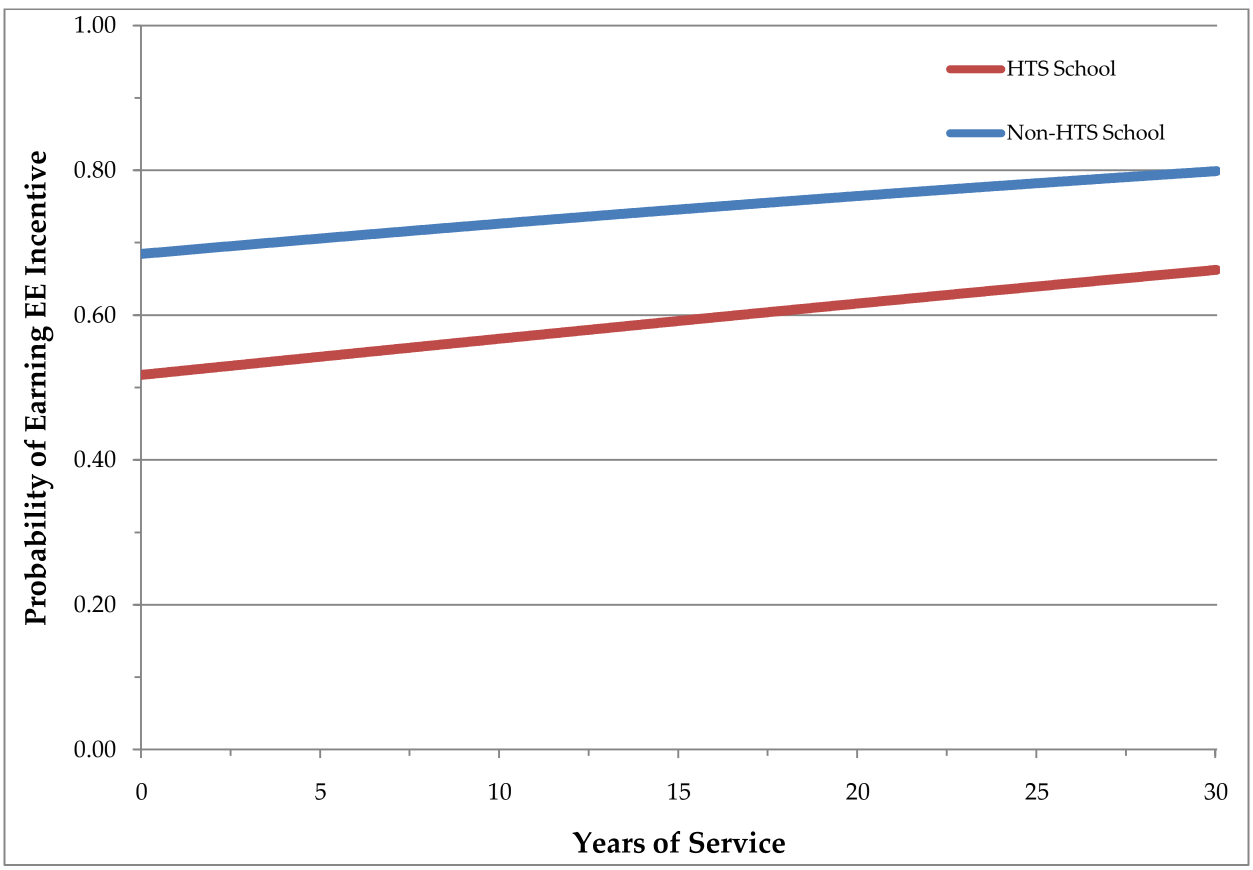
	Math CSAP				Reading CSAP			
	Elementary		Secondary		Elementary		Secondary	
	NonEE	EE	NonEE	EE	NonEE	EE	NonEE	EE
Teacher Years of Service	9.3	9.6	10.9	9.5	8.6	10.0	9.5	8.6
% of Teachers in Hard to Serve Schools	47.7	47.0	41.5	52.8	47.6	39.0	59.6	28.9
% ELL Students in School	23.8	25.2	9.1	13.7	20.0	22.4	14.9	11.4
% GT Students in School	14.8	17.4	22.6	23.7	15.4	18.6	20.1	25.6
% FRL Students in School	70.7	68.1	60.5	66.6	71.0	61.5	69.4	62.7
% Minority Students in School	79.6	74.3	77.0	78.8	79.7	68.0	83.1	77.0
% Teachers Retained in School	87.5	88.6	88.6	87.8	87.4	88.9	89.8	87.2
% FRL Students in Class(es)	72.4	69.8	64.6	65.7	72.7	62.9	69.5	62.1
% Minority Students in Class(es)	84.5	78.5	80.0	76.2	83.3	73.0	82.2	77.0
% "Proficient" in Prior Year - Class	30.6	30.1	18.7	19.9	39.4	43.1	34.2	41.9
% "Advanced" in Prior Year - Class	14.8	15.4	9.7	13.7	2.9	4.1	3.7	6.7
% "Unsatisfactory" in Prior Year - Class	14.0	12.9	26.8	23.1	14.8	14.0	17.2	12.7
% Students Met Growth Criteria	34.6	66.3	37.1	62.7	34.9	62.2	38.2	59.1
Mean CSAP Scale Score	469.8	503.6	535.4	553.7	572.1	600.7	610.5	653.0
Mean CSAP Prior Scale Score	447.7	456.4	526.8	532.5	549.5	565.3	601.9	634.1
Mean CSAP Z-Score	-0.219	0.147	-0.143	0.135	-0.093	0.202	-0.222	0.184
Mean prior CSAP Z-Score	-0.077	-0.039	-0.100	0.018	-0.074	-0.005	-0.243	0.073

Appendix B.
Logistic Regression Analysis of EE Award

Table B-1.
2008-2009 descriptive statistics by CSAP subject, school level, and exceeds expectations incentive receipt

	B	S.E.	Wald	Df	Sig.	Odds Ratio
Years of Service	.020	.014	1.981	1	.159	1.020
Masters +	.244	.534	.208	1	.649	1.276
Hard to Serve	-.705	.244	8.328	1	.004	.494
N of Students	-.007	.005	2.539	1	.111	.993
% Teachers Retained (prior yr)	1.516	.904	2.812	1	.094	4.552
Elementary Dummy	.741	.287	6.692	1	.010	2.098
Constant	-.966	.849	1.294	1	.255	.381

Figure B-1.
Predicted probability of Exceeds Expectations incentive for teachers in hard-to-serve and non-hard-to-serve-schools



Appendix C.

Descriptive Statistics: Exceeds Expectations Achievement Analysis

Table C-1.

Mean student achievement, by student and teacher characteristics

		Math		Reading	
		NonEE	EE	NonEE	EE
Grade 5	Non-HTS				
	Non-FRL	361 (37.1%)	613 (62.9%)	338 (33.6%)	669 (66.4%)
	FRL	509 (48.1%)	549 (51.9%)	424 (41.1%)	607 (58.9%)
	Total	870 (42.8%)	1162 (57.2%)	762 (37.4%)	1276 (62.6%)
	HTS				
	Non-FRL	97 (55.7%)	77 (44.3%)	77 (50.3%)	76 (49.7%)
	FRL	686 (47.4%)	761 (52.6%)	604 (43.9%)	772 (56.1%)
	Total	783 (48.3%)	838 (51.7%)	681 (44.5%)	848 (55.5%)
Grade 8	Non-HTS				
	Non-FRL	868 (84.6%)	158 (15.4%)	301 (43.0%)	399 (57.0%)
	FRL	747 (68.7%)	341 (31.3%)	441 (59.9%)	295 (40.1%)
	Total	1615 (76.4%)	499 (23.6%)	742 (51.7%)	694 (48.3%)
	HTS				
	Non-FRL	273 (69.6%)	119 (30.4%)	206 (70.1%)	88 (29.9%)
	FRL	1239 (65.4%)	656 (34.6%)	1028 (66.3%)	522 (33.7%)
	Total	1512 (66.1%)	775 (33.9%)	1234 (66.9%)	610 (33.1%)
Grade 10	Non-HTS				
	Non-FRL	666 (48.7%)	702 (51.3%)	291 (28.5%)	731 (71.5%)
	FRL	592 (62.3%)	359 (37.7%)	305 (36.6%)	529 (63.4%)
	Total	1258 (54.2%)	1061 (45.8%)	596 (32.1%)	1260 (67.9%)
	HTS				
	Non-FRL	119 (66.9%)	59 (33.1%)	40 (36.4%)	70 (63.6%)
	FRL	538 (63.2%)	313 (36.8%)	185 (40.2%)	275 (59.8%)
	Total	657 (63.8%)	372 (36.2%)	225 (39.5%)	345 (60.5%)

Table C-2.
Mean student achievement, by student and teacher characteristics

		Grade 5		Grade 8		Grade 10	
Subject	Student Group	EE	Non-EE	EE	Non-EE	EE	Non-EE
Mean CSAP Math (SD)	NonFRL-NonHTS	570.1 (69.0)	534.0 (78.1)	577.6 (60.8)	585.5 (58.6)	623.7 (71.3)	573.8 (65.1)
	Non FRL-HTS	495.2 (69.3)	471.4 (67.5)	569.6 (61.2)	549.6 (66.9)	565.3 (53.7)	532.5 (65.2)
	FRL-NonHTS	501.2 (69.0)	472.7 (65.5)	529.0 (67.9)	537.2 (56.6)	553.7 (76.5)	546.8 (60.5)
	FRL-HTS	483.9 (64.0)	468.4 (66.8)	539.4 (55.8)	521.1 (57.6)	555.7 (56.7)	534.1 (59.0)
Prior Year Mean CSAP Math (SD)	NonFRL-NonHTS	531.5 (69.0)	521.3 (83.9)	556.0 (66.3)	578.1 (63.1)	613.4 (63.3)	575.5 (58.4)
	Non FRL-HTS	472.8 (66.6)	452.3 (66.5)	549.0 (70.3)	536.1 (76.2)	563.1 (53.0)	533.7 (60.6)
	FRL-NonHTS	457.8 (72.8)	453.2 (73.3)	509.3 (76.8)	520.9 (62.4)	551.7 (74.0)	546.3 (53.5)
	FRL-HTS	438.0 (65.6)	442.0 (71.9)	503.4 (64.0)	505.5 (60.0)	550.8 (56.0)	532.5 (58.3)
Mean CSAP Reading (SD)	NonFRL-NonHTS	655.3 (55.3)	631.9 (58.0)	695.1 (49.1)	662.3 (49.9)	714.8 (51.6)	673.4 (59.2)
	Non FRL-HTS	589.0 (65.4)	566.9 (63.2)	630.8 (51.4)	639.0 (64.6)	647.1 (56.0)	619.9 (70.7)
	FRL-NonHTS	591.2 (59.3)	572.5 (68.4)	622.6 (65.1)	604.8 (71.5)	668.8 (54.1)	650.2 (54.2)
	FRL-HTS	576.9 (54.9)	563.3 (61.1)	612.5 (53.2)	598.0 (58.1)	647.8 (41.2)	636.2 (57.9)
Prior Year Mean CSAP Reading (SD)	NonFRL-NonHTS	617.1 (50.7)	614.1 (52.1)	685.2 (56.8)	655.9 (53.0)	690.9 (48.2)	660.6 (52.8)
	Non FRL-HTS	567.1 (44.6)	560.2 (48.8)	610.5 (65.2)	631.2 (65.1)	628.0 (54.3)	602.2 (49.6)
	FRL-NonHTS	556.5 (57.3)	557.5 (61.2)	595.2 (86.2)	598.1 (80.1)	647.8 (47.4)	630.5 (47.8)
	FRL-HTS	538.3 (62.7)	539.2 (57.4)	590.3 (63.7)	586.9 (61.6)	615.4 (40.6)	619.0 (46.5)

Appendix D.
Student Achievement Analysis Results – Grade 5

Table D-1.
Final estimation of grade 5 HLM reading achievement fixed effects

Fixed Effect		Coefficient	SE	*t*-ratio	DF	*p*-value
For INTRCPT1, π_0	For INTRCPT2, β_{00}					
	INTRCPT3, γ_{000}	213.4	17.5	12.2	71	<0.001
	HTS, γ_{001}	-51.3	21.6	-2.4	71	0.0
	For EERDG, β_{01}					
	INTRCPT3, γ_{010}	-86.6	20.9	-4.1	95	<0.001
	HTS, γ_{011}	138.7	27.4	5.1	95	<0.001
For RDG_PRIO slope, π_1	For INTRCPT2, β_{10}					
	INTRCPT3, γ_{100}	0.69	0.03	24.3	3419	<0.001
	HTS, γ_{101}	0.06	0.04	1.8	3419	0.1
	For EERDG, β_{11}					
	INTRCPT3, γ_{110}	0.16	0.03	4.8	3419	<0.001
	HTS, γ_{111}	-0.24	0.05	-5.3	3419	<0.001
For FRL slope, π_2	For INTRCPT2, β_{20}					
	INTRCPT3, γ_{200}	-15.7	3.4	-4.6	3419	<0.001
	HTS, γ_{201}	11.0	5.1	2.2	3419	0.0
	For EERDG, β_{21}					
	INTRCPT3, γ_{210}	7.5	4.2	1.8	3419	0.1
	HTS, γ_{211}	-6.0	7.3	-0.8	3419	0.4

Figure D-1.
Grade 5 predicted reading scores by student type

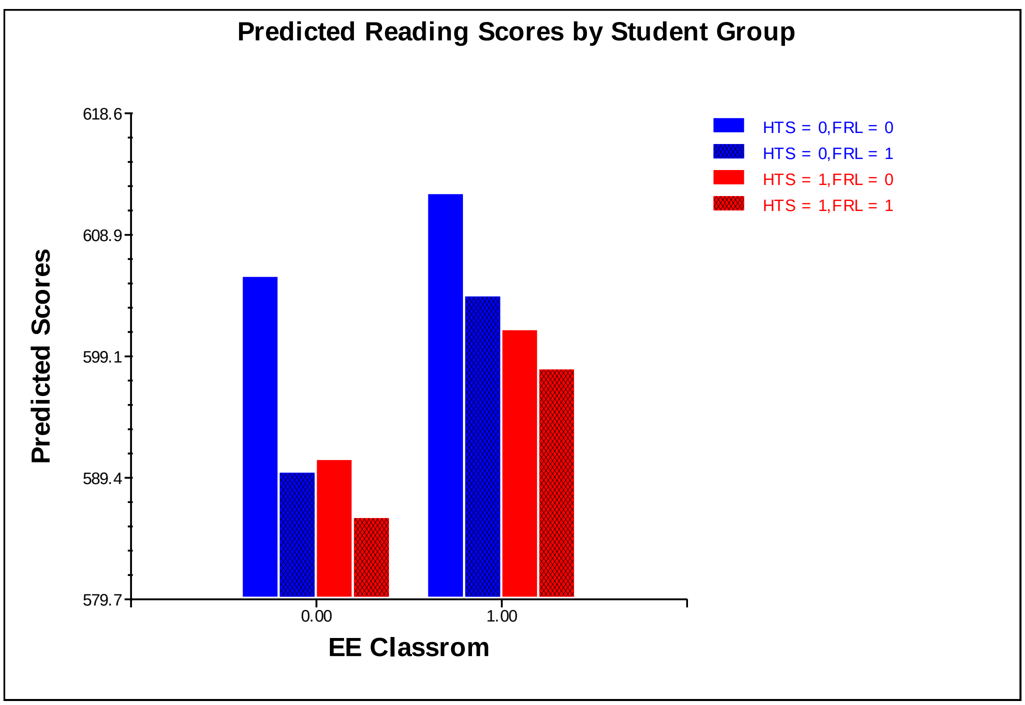


Figure D-2.

Grade 5 predicted reading score plot for EE & FRL, conditioned on prior achievement

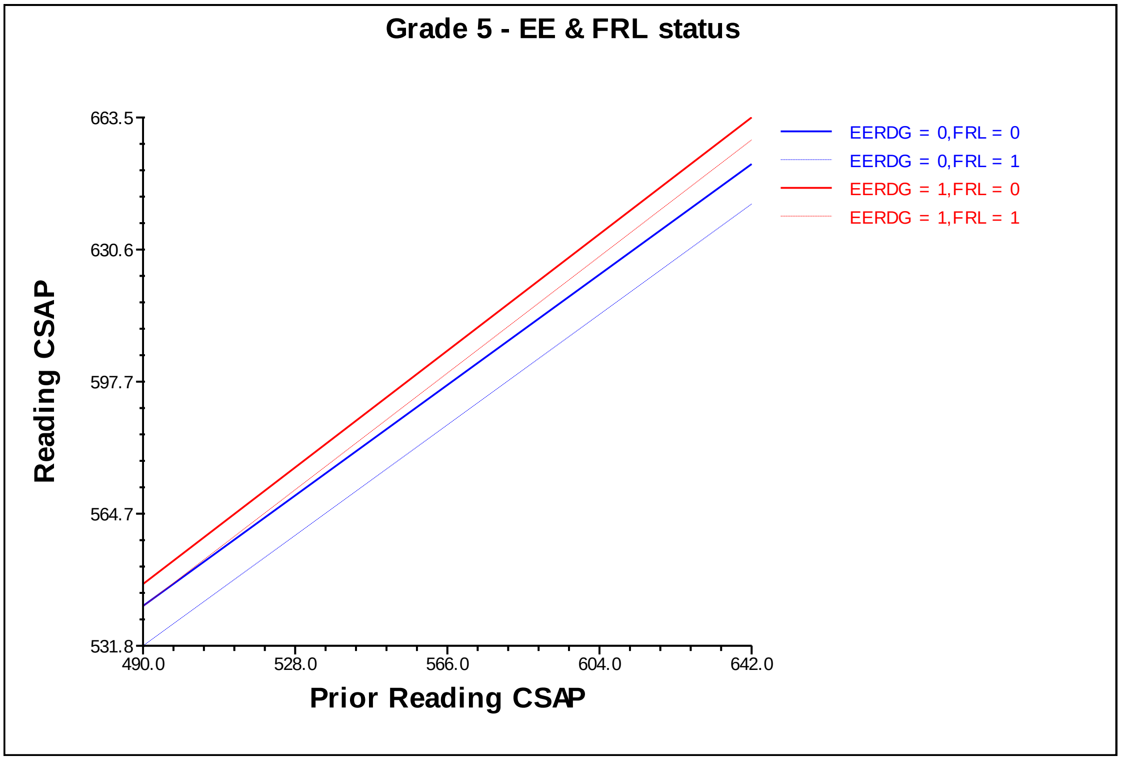


Figure D-3.

Grade 5 predicted reading score plot for EE & HTS, conditioned on prior achievement

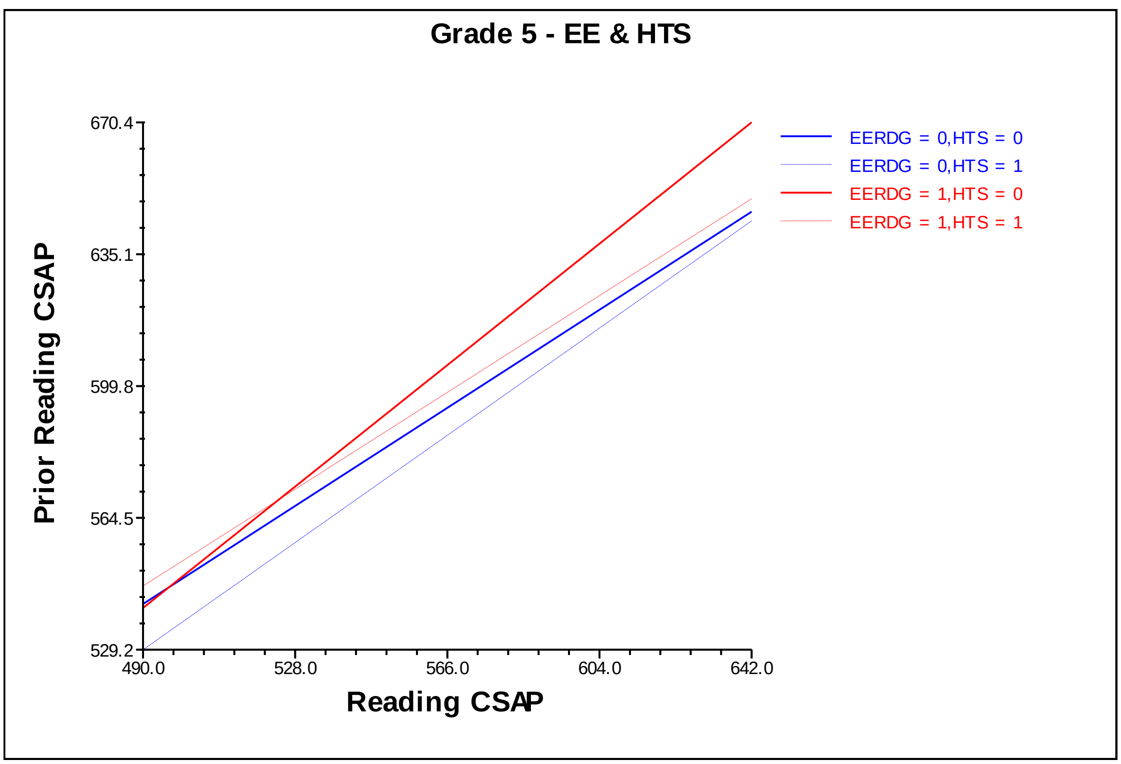


Table D-2.

Final estimation of grade 5 HLM math achievement fixed effects

Fixed Effect		Coefficient	SE	*t*-ratio	DF	*p*-value
For INTRCPT1, π_0	For INTRCPT2, β_{00}					
	INTRCPT3, γ_{000}	148.1	9.8	15.1	73	<0.001
	HTS, γ_{001}	-41.5	13.1	-3.2	73	0.0
	For EEMATH, β_{01}					
	INTRCPT3, γ_{010}	-10.9	13.1	-0.8	101	0.4
	HTS, γ_{011}	32.6	18.7	1.7	101	0.1
For MATH_PRI slope, π_1	For INTRCPT2, β_{10}					
	INTRCPT3, γ_{100}	0.76	0.02	41.9	3415	<0.001
	HTS, γ_{101}	0.07	0.02	2.7	3415	0.0
	For EEMATH, β_{11}					
	INTRCPT3, γ_{110}	0.06	0.02	2.3	3415	0.0
	HTS, γ_{111}	-0.06	0.04	-1.7	3415	0.1
For FRL slope, π_2	For INTRCPT2, β_{20}					
	INTRCPT3, γ_{200}	-14.4	3.1	-4.6	3415	<0.001
	HTS, γ_{201}	12.9	5.1	2.6	3415	0.0
	For EEMATH, β_{21}					
	INTRCPT3, γ_{210}	11.0	4.1	2.7	3415	0.0
	HTS, γ_{211}	-13.3	7.2	-1.8	3415	0.1

Figure D-4.

Grade 5 predicted math scores by student type

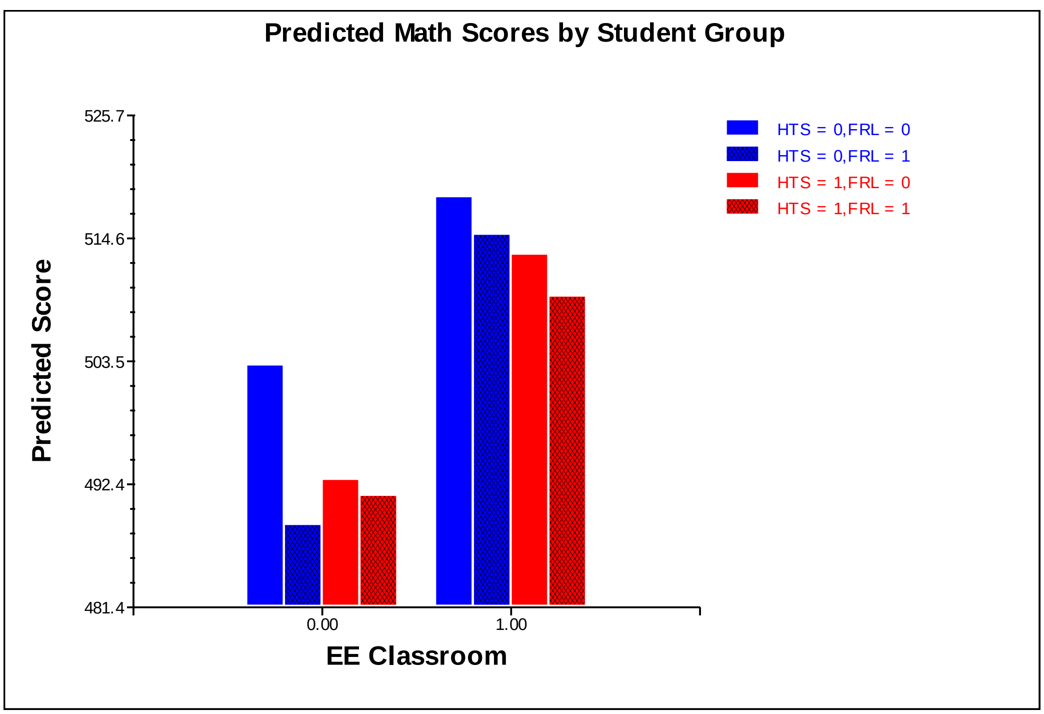


Figure D-5.
Grade 5 predicted math score plot for EE & FRL, conditioned on prior achievement

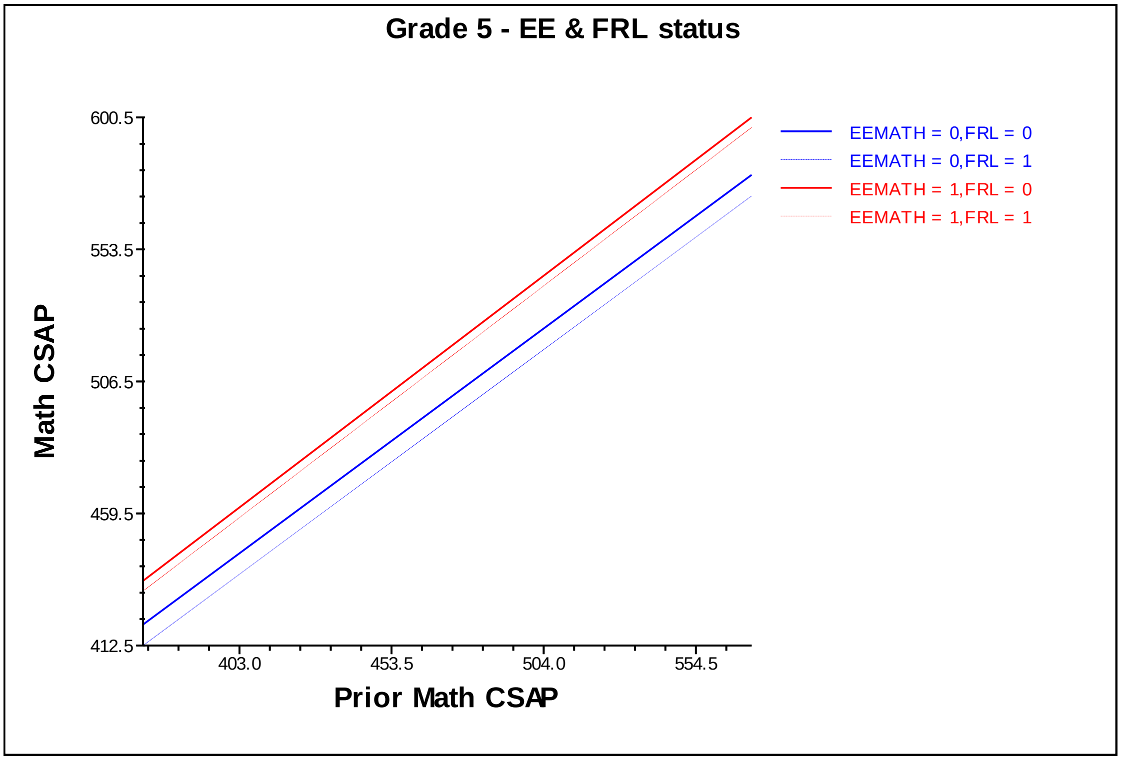
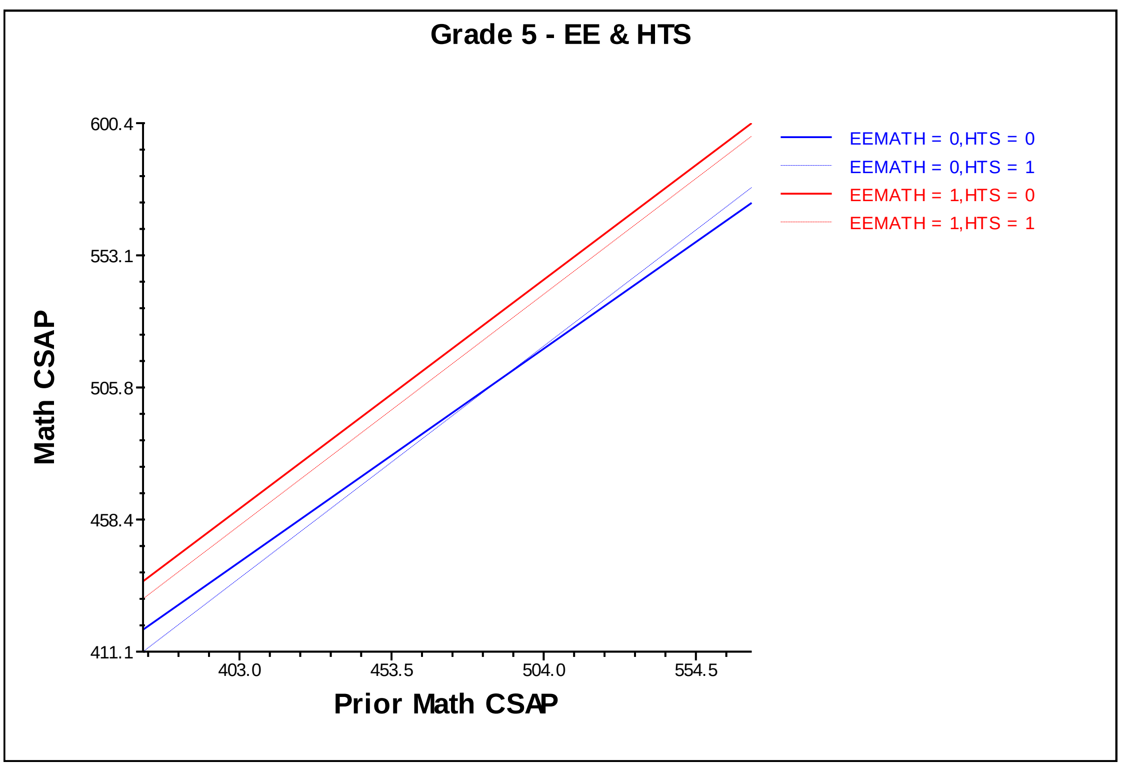


Figure D-6.
Grade 5 predicted reading score plot for EE & HTS, conditioned on prior achievement



Appendix E.
Student Achievement Analysis Results – Grade 8

Table E-1.
Final estimation of grade 8 HLM reading achievement fixed effects

Fixed Effect			Coefficient	SE	*t*-ratio	DF	*p*-value
For INTRCPT1, π_0	For INTRCPT2, β_{00}						
		INTRCPT3, γ_{000}	172.5	17.1	10.1	31	<0.001
		HTS, γ_{001}	54.2	20.2	2.7	31	0.01
	For EERDG, β_{01}						
		INTRCPT3, γ_{010}	24.9	2.3	56.0	0	<0.001
		HTS, γ_{011}	31.4	-1.1	56.0	0	<0.001
For RDG_PRIO slope, π_1	For INTRCPT2, β_{10}						
		INTRCPT3, γ_{100}	0.73	0.02	29.7	3250	<0.001
		HTS, γ_{101}	-0.09	0.03	-3.2	3250	0.00
	For EERDG, β_{11}						
		INTRCPT3, γ_{110}	-0.08	0.04	-2.1	3250	0.03
		HTS, γ_{111}	0.06	0.05	1.3	3250	0.18
For FRL slope, π_2	For INTRCPT2, β_{20}						
		INTRCPT3, γ_{200}	-8.1	2.9	-2.8	3250	0.01
		HTS, γ_{201}	-1.0	3.8	-0.3	3250	0.79
	For EERDG, β_{21}						
		INTRCPT3, γ_{210}	3.0	4.7	0.6	3250	0.53
		HTS, γ_{211}	4.2	7.0	0.6	3250	0.55

Figure E-1.
Grade 8 predicted reading scores by student type

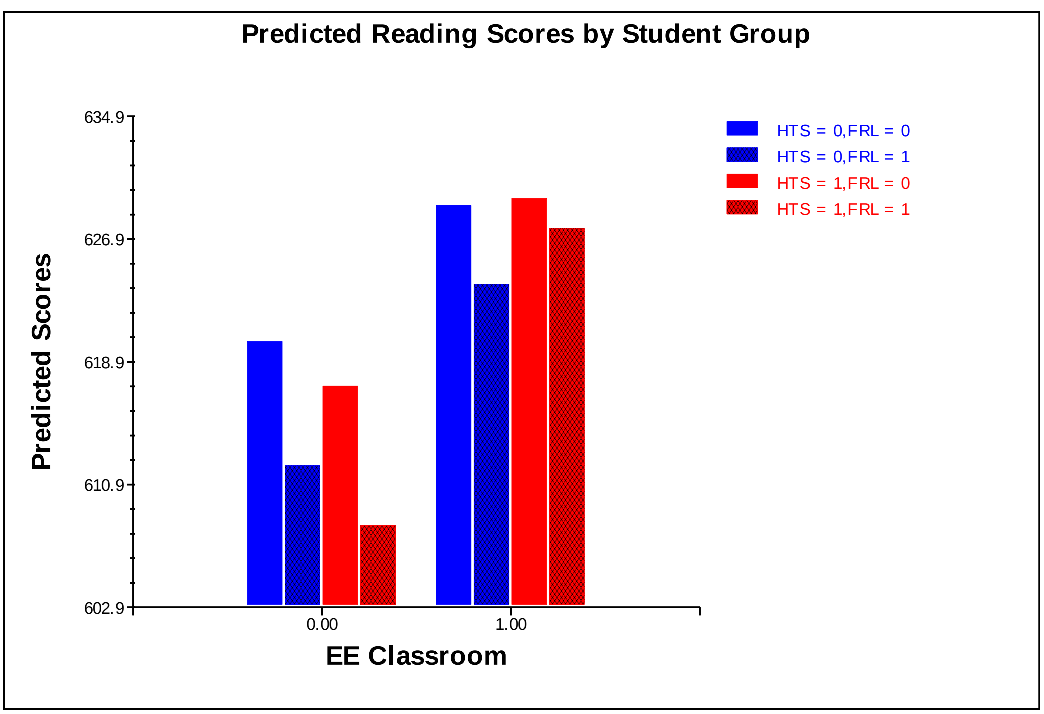


Figure E-2.
Grade 8 predicted reading score plot for EE & FRL, conditioned on prior achievement

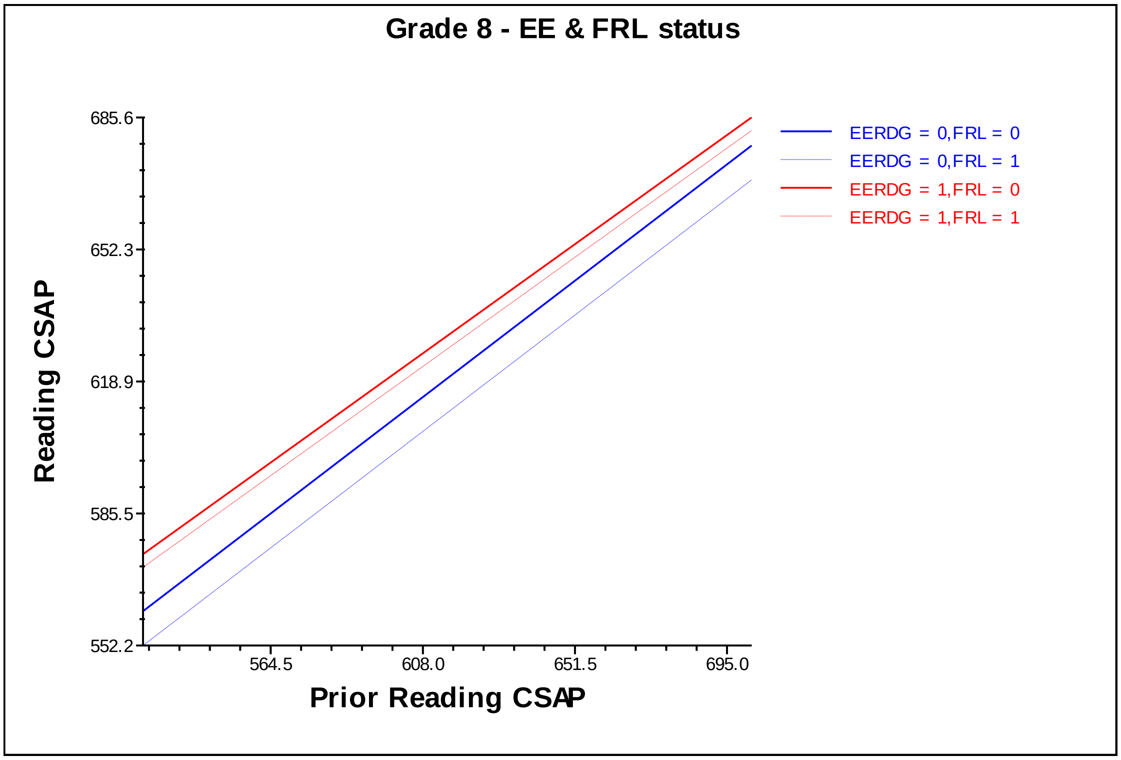


Figure E-3.
Grade 8 predicted reading score plot for EE & HTS, conditioned on prior achievement

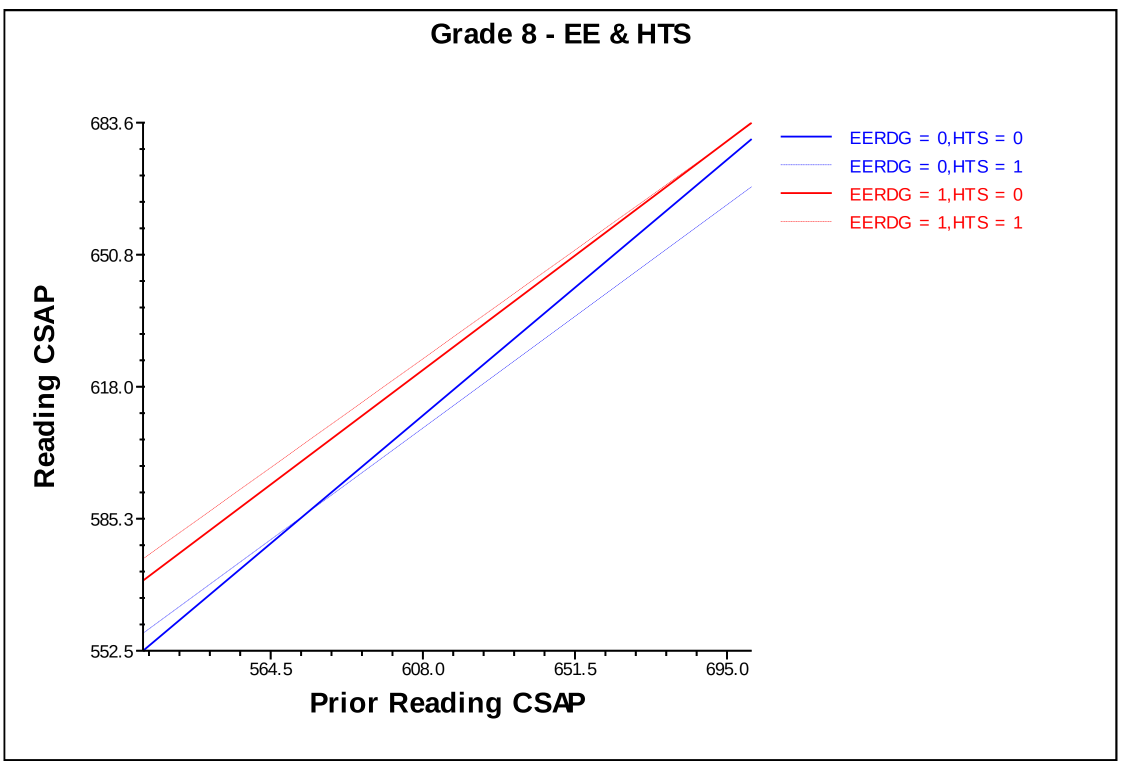


Table E-2.
Final estimation of grade 8 HLM math achievement fixed effects

Fixed Effect		Coefficient	SE	*t*-ratio	DF	*p*-value
For INTRCPT1, π_0	For INTRCPT2, β_{00}					
	INTRCPT3, γ_{000}	147.3	9.7	15.1	37	<0.001
	HTS, γ_{001}	-24.7	12.6	-2.0	37	0.06
	For EEMATH, β_{01}					
	INTRCPT3, γ_{010}	22.8	17.4	1.3	57	0.20
	HTS, γ_{011}	17.4	22.0	0.8	57	0.43
For MATH_PRI slope, π_1	For INTRCPT2, β_{10}					
	INTRCPT3, γ_{100}	0.76	0.02	46.1	4256	<0.001
	HTS, γ_{101}	0.04	0.02	2.0	4256	0.04
	For EEMATH, β_{11}					
	INTRCPT3, γ_{110}	-0.02	0.03	-0.8	4256	0.41
	HTS, γ_{111}	-0.03	0.04	-0.9	4256	0.37
For FRL slope, π_2	For INTRCPT2, β_{20}					
	INTRCPT3, γ_{200}	-2.1	2.0	-1.0	4256	0.30
	HTS, γ_{201}	-2.7	2.9	-0.9	4256	0.35
	For EEMATH, β_{21}					
	INTRCPT3, γ_{210}	-6.2	4.4	-1.4	4256	0.16
	HTS, γ_{211}	10.6	6.1	1.7	4256	0.09

Figure E-4.
Grade 8 predicted math scores by student type

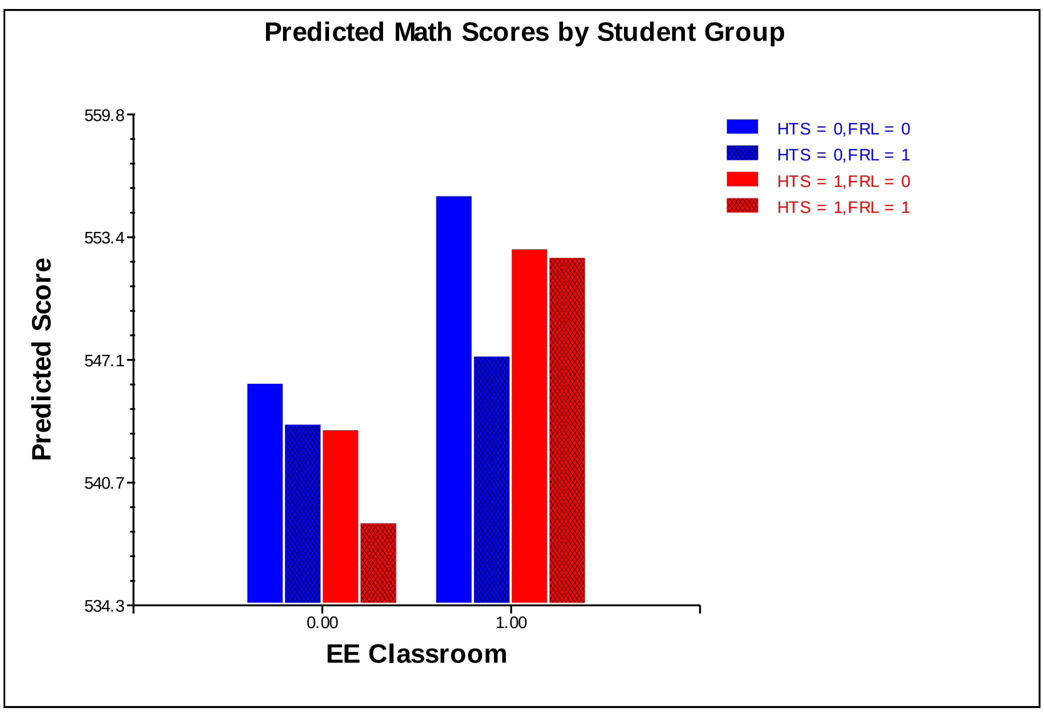


Figure E-5.

Grade 8 predicted math score plot for EE & FRL, conditioned on prior achievement

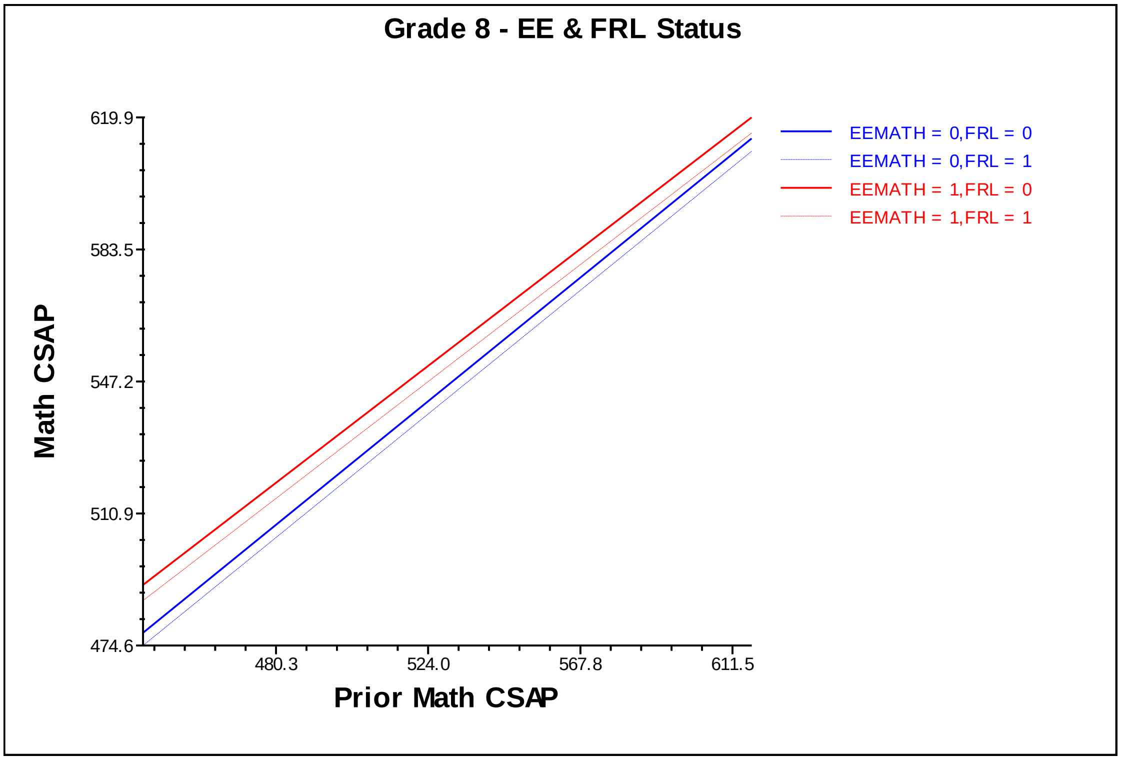
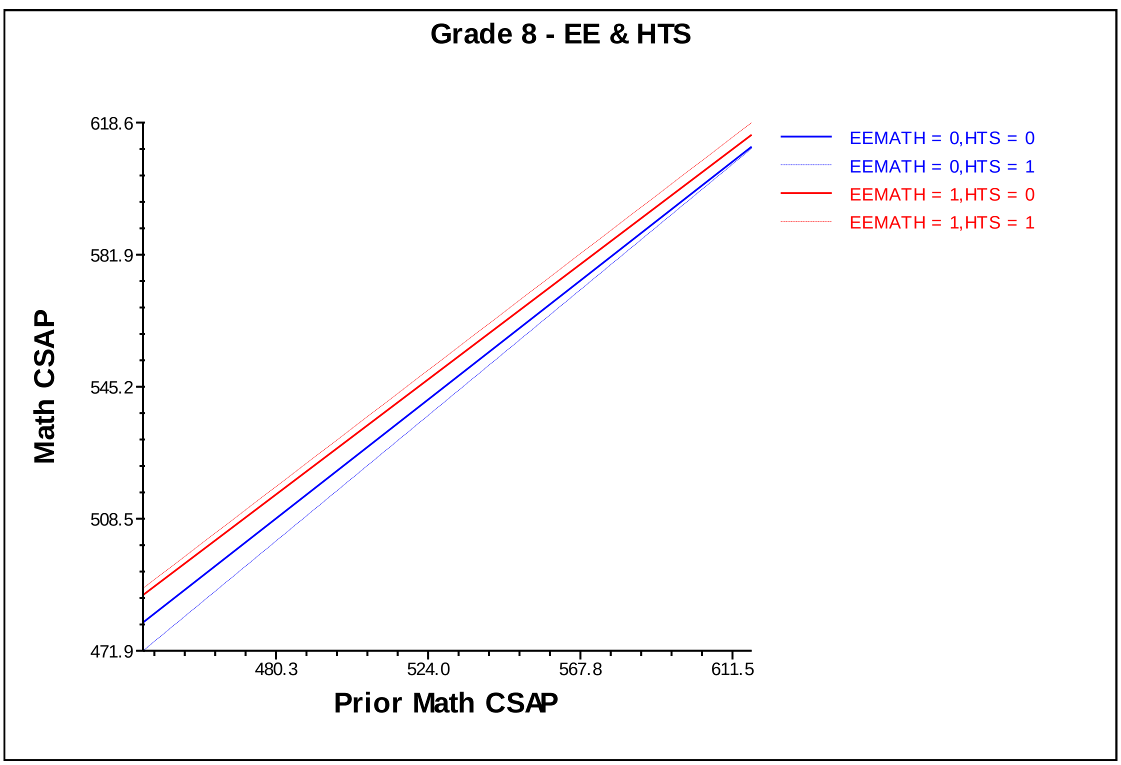


Figure E-6.

Grade 8 predicted reading score plot for EE & HTS, conditioned on prior achievement



Appendix F.
Student Achievement Analysis Results – Grade 10

Table F-1.
Final estimation of grade 10 HLM reading achievement fixed effects

Fixed Effect		Coefficient	SE	*t*-ratio	DF	*p*-value
For INTRCPT1, π_0	For INTRCPT2, β_{00}					
	INTRCPT3, γ_{000}	173.7	21.6	8.0	25	<0.001
	HTS, γ_{001}	-189.7	35.5	-5.4	25	<0.001
	For EERDG, β_{01}					
	INTRCPT3, γ_{010}	-44.6	26.2	-1.7	56	0.10
	HTS, γ_{011}	228.3	47.5	4.8	56	<0.001
For RDG_PRIO slope, π_1	For INTRCPT2, β_{10}					
	INTRCPT3, γ_{100}	0.75	0.03	22.8	2306	<0.001
	HTS, γ_{101}	0.29	0.06	5.2	2306	<0.001
	For EERDG, β_{11}					
	INTRCPT3, γ_{110}	0.09	0.04	2.4	2306	0.02
	HTS, γ_{111}	-0.37	0.07	-5.0	2306	<0.001
For FRL slope, π_2	For INTRCPT2, β_{20}					
	INTRCPT3, γ_{200}	-2.5	3.2	-0.8	2306	0.43
	HTS, γ_{201}	13.3	6.4	2.1	2306	0.04
	For EERDG, β_{21}					
	INTRCPT3, γ_{210}	-0.4	3.8	-0.1	2306	0.92
	HTS, γ_{211}	-2.9	8.0	-0.4	2306	0.72

Figure F-1.
Grade 10 predicted reading scores by student type

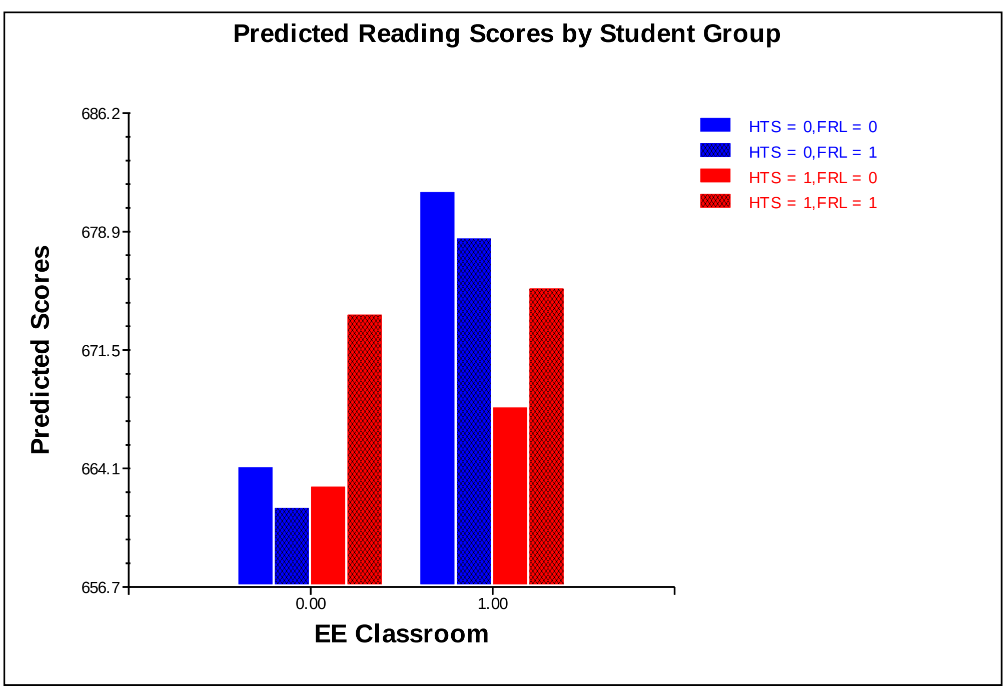


Figure F-2.
Grade 10 predicted reading score plot for EE & FRL, conditioned on prior achievement

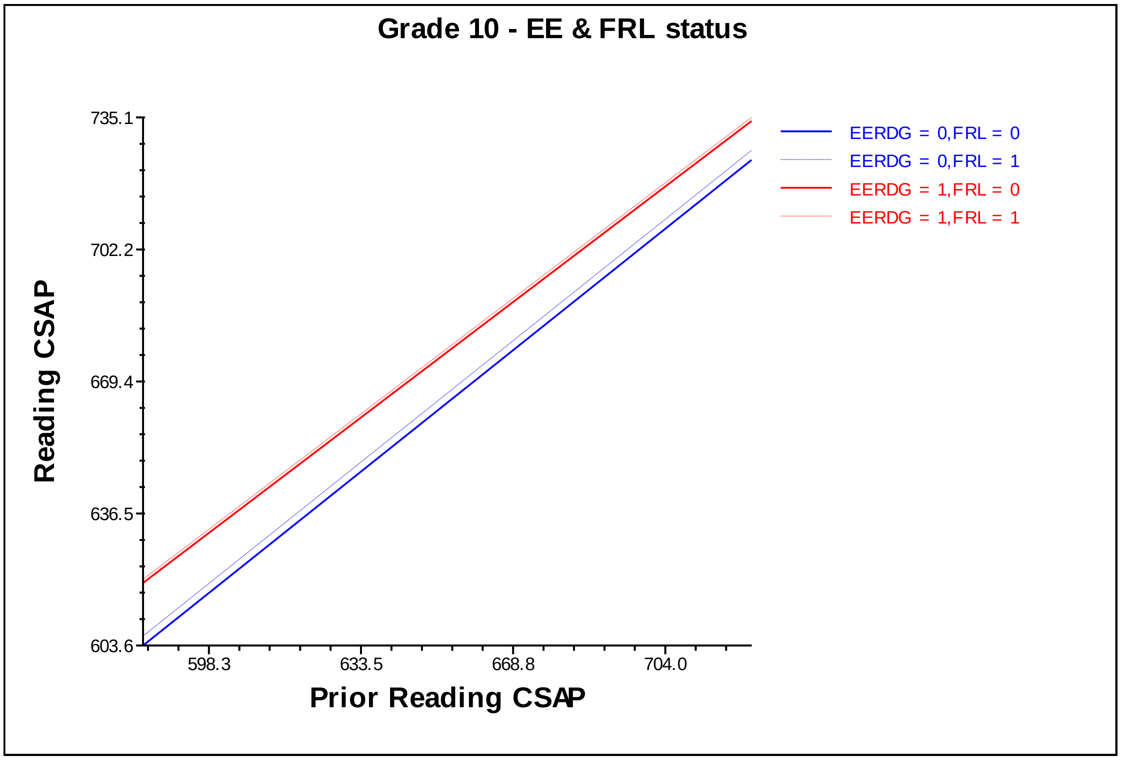


Figure E-3.
Grade 10 predicted reading score plot for EE & HTS, conditioned on prior achievement

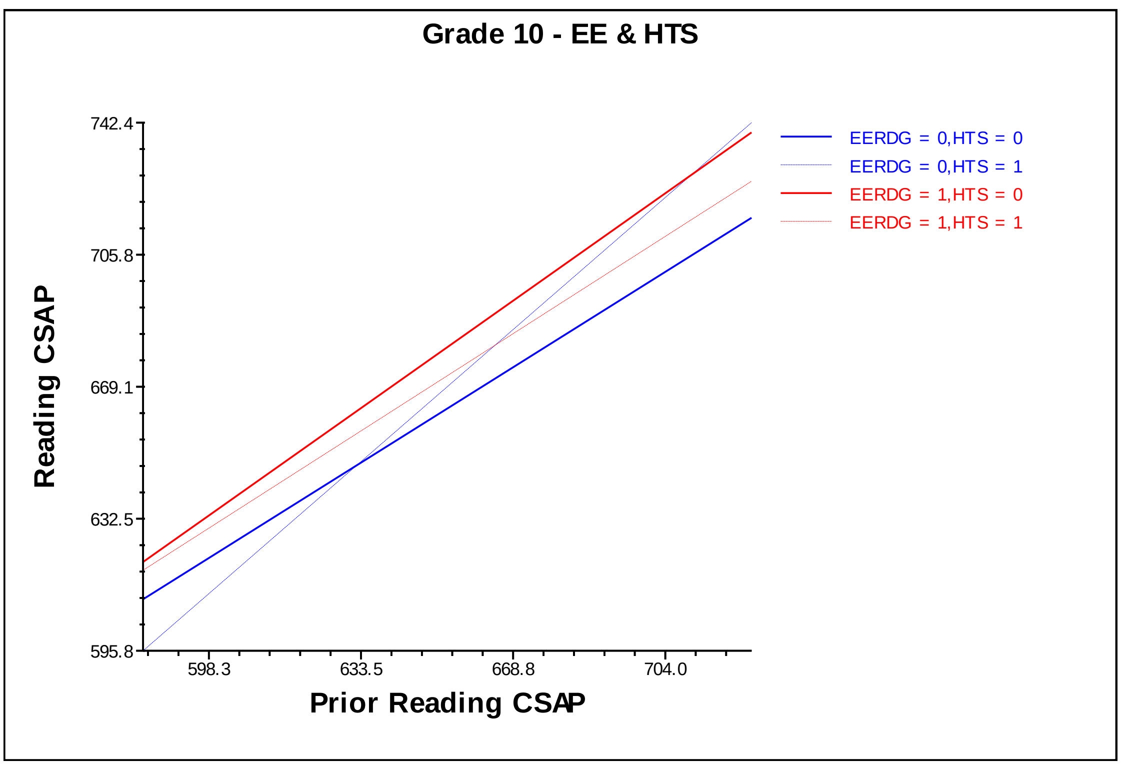


Table F-2.
Final estimation of grade 10 HLM math achievement fixed effects

Fixed Effect		Coefficient	SE	t-ratio	DF	p-value
For INTRCPT1, π_0	For INTRCPT2, β_{00}					
	INTRCPT3, γ_{000}	79.8	11.5	7.0	22	<0.001
	HTS, γ_{001}	33.7	18.7	1.8	22	0.09
	For EEMATH, β_{01}					
	INTRCPT3, γ_{010}	5.9	16.3	0.4	69	0.72
	HTS, γ_{011}	-12.1	29.1	-0.4	69	0.68
For MATH_PRI slope, π_1	For INTRCPT2, β_{10}					
	INTRCPT3, γ_{100}	0.87	0.02	44.1	3199	<0.001
	HTS, γ_{101}	-0.08	0.03	-2.4	3199	0.02
	For EEMATH, β_{11}					
	INTRCPT3, γ_{110}	0.01	0.03	0.3	3199	0.76
	HTS, γ_{111}	0.02	0.05	0.3	3199	0.74
For FRL slope, π_2	For INTRCPT2, β_{20}					
	INTRCPT3, γ_{200}	-3.9	2.3	-1.7	3199	0.09
	HTS, γ_{201}	9.0	4.5	2.0	3199	0.05
	For EEMATH, β_{21}					
	INTRCPT3, γ_{210}	-5.3	3.4	-1.5	22	0.14
	HTS, γ_{211}	1.0	7.0	0.1	22	0.89

Figure F-4.
Grade 10 predicted math scores by student type

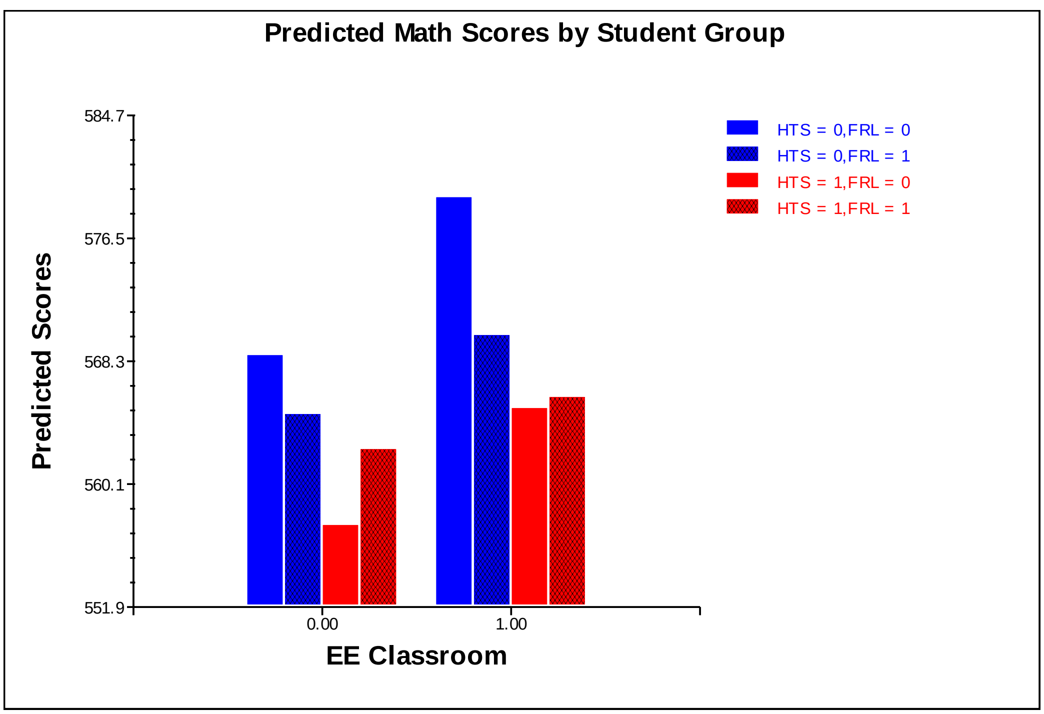


Figure F-5.
Grade 10 predicted math score plot for EE & FRL, conditioned on prior achievement

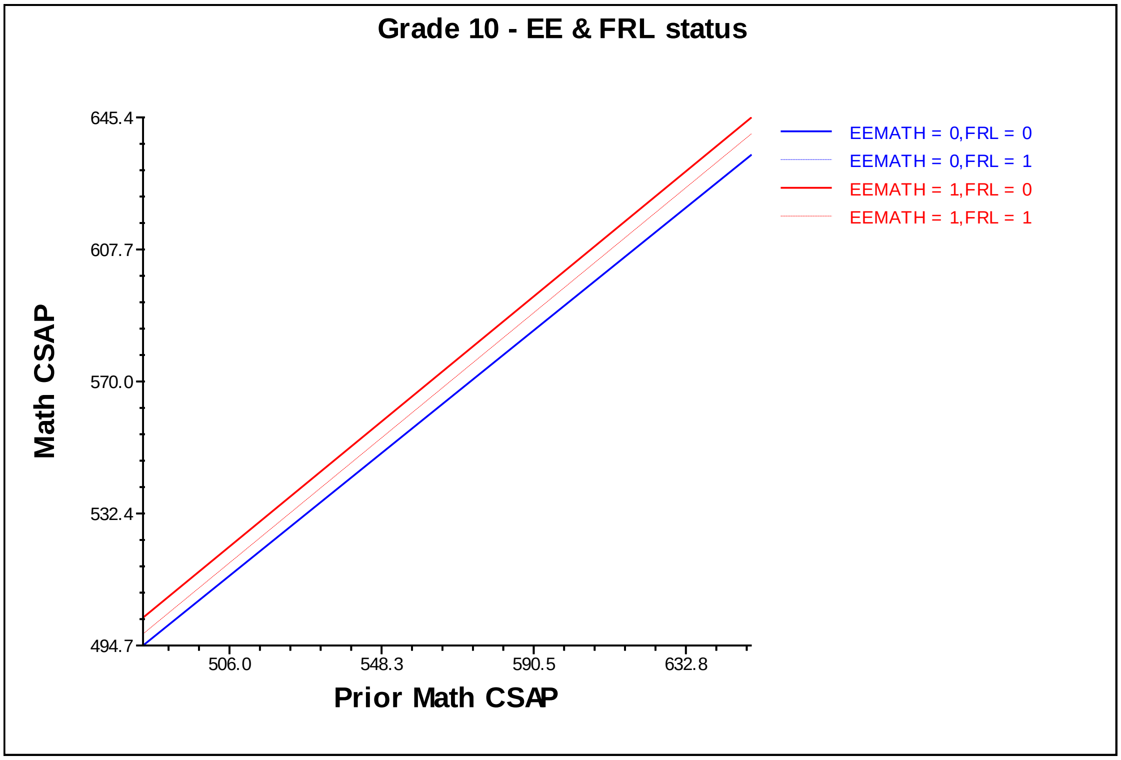


Figure F-6.
Grade 10 predicted reading score plot for EE & HTS, conditioned on prior achievement