Assessing the Effectiveness of Tennessee's Pre-Kindergarten Program: Second Interim Report

August 18, 2008

995 Goodale Blvd ● Columbus, Ohio 43212

Phone: (614) 220-8860 ● Fax: (614) 220-8845
www.strategicresearchgroup.com

The **Strategic Research Group (SRG)** is a full-service research firm that provides data collection, consultative, and research services. SRG specializes in conducting public opinion surveys, program evaluations, policy assessments, customer satisfaction studies, and community needs assessments on national, state, and local levels.

Foreword and Acknowledgements

This **Second Interim Report**, produced under contract with the Tennessee State Comptroller's Office, provides the results of an analysis of student outcomes for students who participated in Tennessee's Pre-K program between 1998-1999 and 2005-2006. Statistical analyses explored the short- and long-term impact of Pre-K participation on student assessments in Kindergarten through Fifth Grade. This report also provides a descriptive overview of Pre-Kindergarten students and programs in Tennessee (see Appendix A) and offers a detailed overview of the research methodology used to assess the effectiveness of Pre-Kindergarten programs in Tennessee.

This report has been made possible due to the support from the Tennessee Department of Education (TDOE) and the Tennessee State Comptroller's Office of Education Accountability. In particular, we wish to thank the following individuals for their expertise and assistance in obtaining the data necessary for analysis, providing important background information, and for their ongoing support of this research:

Connie Casha, Director of Early Childhood Programs for the Office of Early Learning

Corey Chatis, Director of Data Quality

Dr. Phillip Doss, Director of the Tennessee State Comptroller's Office of Education Accountability

Dan Long, Senior Executive Director for the TDOE Office of Assessment, Evaluation and Research

Bobbi Lussier, Executive Director for the Office of Early Learning

Tracey Ray, Director of Data Services for the Office of Early Learning

TABLE OF CONTENTS

Executive Summary **5**

Evaluating the Effectiveness of Tennessee's Pre-Kindergarten Program **9**

Background and Objectives of the Present Study 10

Research Design **10**

Methodology **13**

Data Sources 13

1. Pre-Kindergarten Demographic File 13

2. Education Information System Data 14

3. K-12 Student Assessment Data 15

Data Management 17

Sampling Strategy 22

Analytic Approach 25

Results **34**

Next Steps and Subsequent Reports **57**

Appendix A. Background and Implementation of Tennessee's Pre-Kindergarten Program **59**

Pre-K Pilot Sites 60

Student Eligibility 61

Classroom Requirements 61

Program Effectiveness 62

Appendix B. Technical Specification of Models **64**

Executive Summary

The State of Tennessee has commissioned an evaluation of the effectiveness of its Pre-Kindergarten (Pre-K)[1] program through a secondary data analysis (i.e., analysis of existing data) of student outcomes comparing Pre-K participants to a comparison group of students who did not attend state-funded Pre-K. The primary objective of the project is to assess whether children who attended a Tennessee-funded Pre-K program perform better academically in the short and long term than a comparable group of peers who did not attend Tennessee's Pre-K program. The evaluation will also investigate whether various characteristics of Tennessee's Pre-K programs impact short- and long-term achievement among students who attended these programs.[2]

The objective of this Second Interim Report is to build on the results presented in the First Interim Report in December 2007 and to provide the Office of Education Accountability (OEA) and the Tennessee General Assembly a report on the results of an analysis of student outcomes for students who participated in Tennessee's Pre-K program between 1998-1999 and 2005-2006. As specified by Request for Proposal (RFP) 308.14-004, this Second Interim Report analyzes student achievement as measured by results of standardized tests administered in three academic years: 2004-2005, 2005-2006, and 2006-2007.

For the Second Interim Report, Strategic Research Group (SRG) drew from three data sources: 1) Pre-Kindergarten demographic data, 2) K-12 student assessment data, and 3) Education Information System (EIS) student data from the 2005-2006 and 2006-2007 school years. Pre-K students were identified in assessment records and individually matched to another student with the same demographic characteristics in the same school and/or district who did not attend Pre-K. This rigorous precision matching technique was employed to construct a random sample of non-Pre-K students that matched the Pre-K group as closely as possible in all possible respects given the data available for analysis.

The First Interim Report (November, 2007) analyzed student assessment data between 1999-2000 and 2003-2004. Due to small sample sizes and some missing data in these early years of the program, separate analyses were conducted for each grade level each year. The analytic approach taken in this Second Interim Report differed from the approach taken in the First Interim Report given a larger number students had participated in Pre-K in the timeframe under study and there was an opportunity for longitudinal analysis, or an analysis of student outcomes over time. Data were analyzed using random effects models, also referred to as hierarchical linear models or multilevel models.

---

[1] Throughout this report, the term "Pre-Kindergarten and its abbreviation "Pre-K" are used to refer specifically to Tennessee's state-funded Pre-Kindergarten program and not any other type of early childhood education program. The term "non-Pre-K" is used to refer to students who did not attend Tennessee's Pre-K program, although they may have participated in other early childhood education programs.

[2] The present report explores student-level characteristics as predictors of student outcomes, including race, gender, Free/Reduced Price Lunch status, and other characteristics. Analysis of program-level characteristics as predictors of student outcomes depends on the availability of such data from TDOE. More discussion with the Office of the Comptroller and the Office of Early Learning is necessary to identify appropriate variables to include in a program-level analysis as well as to identify a data sources for these variables. If possible, such analysis will be included in a subsequent report.

All models evaluated for this report include children's Free/Reduced Price Lunch (FRPL) history and participation in Tennessee state-funded Pre-K as predictors of academic achievement, and the results focus on these two child-level characteristics as well as their interaction. The models were structured in this way for theoretical and practical reasons. The effect of Pre-K participation is of primary interest in this evaluation, and thus that is the central focus of all the analyses. Further, because the Pre-K program specifically targets children deemed "at-risk" and FRPL status is the only consistent variable available for analysis that serves as a proxy for "risk," FRPL status was considered an important variable to include to help interpret any effects associated with Pre-K participation. In addition to these two important variables, exploratory analyses conducted in the First Interim Report did find some effects associated with students' gender and race; thus, all statistical models examined in this study controlled for a child's race and gender. In addition, the models employed in this report also include additional control variables: whether or not a child received special education services within the observed grades, whether or not a child was retained within the observed grades, the average number of days a child was absent from class during the observed timeframe, and whether or not English is the child's primary or native language. These control variables (and their theoretically or statistically relevant interactions) were included to ensure an accurate representation of the population under study and to ensure potentially mitigating effects were accounted for in the models to control for any potential bias.

Short-term Effects of Pre-K Participation (Kindergarten through Second Grade)

Because different assessments are administered at different points in time between Kindergarten and Second Grade, different kinds of statistical models were employed to examine change over time. *Growth curve models* were used to examine change in assessment scores over three time points (for example, Kindergarten through Second Grade), and *difference score models* were used to examine change in assessment scores over two time points (for example, First and Second Grades). Single time point models were used to examine differences between the Pre-K and non-Pre-K groups when an assessment was administered in only one grade. A consistent pattern of results was observed across the assessments administered in Grades K-2, reflecting the short-term effects of Pre-K participation.

For those assessments administered in Kindergarten, Pre-K students scored better than a matched sample of non-Pre-K students in Reading, Language Arts, and Mathematics. There was also a significant difference depending on whether or not a student received FRPL—an indicator of student socioeconomic status. More specifically, among students who did not receive FRPL, those who participated in Pre-K scored higher, on average, than students who did not participate in Pre-K. Further, among students who did receive FRPL, students who attended Pre-K scored higher than students who did not attend Pre-K. While overall students who did not receive FRPL scored higher than students who did receive FRPL, students receiving FRPL who attended Pre-K scored similarly to non-FRPL students who did not attend Pre-K.

Thus, for Reading, Language Arts, and Mathematics, students who participated in Pre-K tended to score significantly higher, on average, on their Kindergarten assessments relative to peers who did not attend Pre-K. Although students' socioeconomic status, as indicated by receipt of FRPL, also had a significant effect, Pre-K participation was associated with higher scores on Kindergarten assessments for students who received FRPL as well as those who did not.

For both types of models, the pattern observed in Kindergarten changed in First and Second Grades. First, the growth curve models which identified an initial difference in Kindergarten assessment scores showed a pattern of convergence over time. In other words, although Pre-K students initially demonstrated an advantage on these assessments over peers who did not

participate in Pre-K, by the Second Grade there was no statistically significant difference in these groups attributable to Pre-K participation. Rather, the Pre-K and non-Pre-K groups appear to converge, although a significant difference was still found for students' FRPL status. That is, by the Second Grade, there was still a statistically significant effect for students' FRPL status such that students who received FRPL tended to score lower, on average, than students who did not. Within these two groups (students who received FRPL and students who did not receive FRPL), students who attended Pre-K scored similarly to students who did not attend Pre-K.

Separate analyses were conducted for assessments administered only in First and Second Grades. Difference score models for these assessments examined whether Pre-K students differed from non-Pre-K students, and whether there was a difference in outcomes due to students' FRPL status in the First and Second Grades. The results of these models indicated that Pre-K participation was not associated with a significant difference in student outcomes although FRPL status was. That is, among students who received FRPL, Pre-K students scored similarly to non-Pre-K students in First and Second Grade. The same pattern was observed for students who did not receive FRPL.

Long-term Effects of Pre-K Participation

As observed in the results of assessments administered in Grades K-2, a consistent pattern also emerged across assessments administered in Grades 3-5. Growth curve models showed only one statistically significant difference associated with Pre-K participation, such that in the Third Grade students who had participated in Pre-K performed slightly better on the Mathematics assessment relative to non-Pre-K students. There were no other significant effects associated with Pre-K participation in Grades 3-5. The difference between students who received FRPL and those who did not (i.e., student socioeconomic status), was associated with a more consistent significant difference in student outcomes across all assessments in Grades 3-5.

Additional Effects

Analysis of fixed effects in the models examined the unique effects of student gender, race, absences, special education, retention, and language (i.e., whether the student is a native English speaker) on student outcomes, and the interactions of these variables with Pre-K participation. In other words, the models analyzed also tested whether there were statistically significant differences in student outcomes associated with these variables and Pre-K participation. All these variables and their interactions were explored for all models (one, two, and three time points) and for all assessments. There were no consistent, statistically significant interactions between Pre-K participation and gender, race, or any of the other predictor variables. This can be interpreted to indicate that overall, the general pattern of short- and long-term results discussed above holds for all subgroups of students—in other words, the general pattern observed is the same for both male and female students, white and non-white students, and so on.

General Summary

On the whole, the results demonstrate an initial advantage associated with Pre-K participation in Kindergarten, and this was the case for students who received FRPL as well as those who did not. However, in analyses of assessments conducted in First and Second Grades, this initial difference was followed by a consistent pattern of convergence. Pre-K participation was associated with significant differences in Kindergarten assessments of Reading, Language Arts, and Mathematics, although students' socioeconomic status (i.e., whether they receive FRPL in the time period under study) also plays a significant role in their outcomes on these

assessments. The models employed in the present study did not find that this relative advantage persisted over time, and as students moved through higher grades, their scores tended to converge so that students receiving FRPL tended to be more similar to one another (irrespective of their participation in Pre-K) and students who did not receive FRPL tended to be more similar to one another (again, irrespective of Pre-K participation).

Additional analysis is warranted to explore these results further, as discussed later in this report. In addition, we wish to address the fact that the statistical control variables used to examine short- and long-term effects associated with Pre-K participation in the present study (retention, attendance, and special education measured after the Pre-K year) are variables which themselves could have been affected by participation in the Pre-K program. These variables were included in the present model due to their theoretical significance. In deciding whether or not to include these variables in the models we did take into consideration the fact that including these controls may have the additional effect of "controlling out" some of the Pre-K program effect. However, failing to account for these important potential sources of variability in students' scores could have led to inaccurate (i.e., biased) results, thereby hampering our ability to provide an accurate reflection of student progress—regardless of Pre-K experience. The possible relationship between Pre-K experience and predictors of academic success such as retention and attendance has not gone unnoticed. However, further exploration is required before a definitive picture of this relationship can be presented. This will be examined in future reports.

Although this report is able to answer with some confidence the primary research questions of interest for the present evaluation, there are many outstanding questions about the impact and effectiveness of the Pre-K program that remain unanswered—and indeed, this report may generate some new questions. We wish to remind the reader that no single study can address every possible question about a program as large as Tennessee's Pre-K program (and one with many passionate proponents), particularly when the data are limited and the methodology is retrospective. However, the present study has employed the best possible methods for the data available to provide as accurate a picture as possible of the performance of Pre-K students, and future reports are planned that will continue to explore these questions as the evaluation progresses.

Evaluating the Effectiveness of Tennessee's Pre-Kindergarten Program

The State of Tennessee has commissioned an evaluation of the effectiveness of its Pre-Kindergarten (Pre-K)[3] program through a secondary data analysis of student outcomes comparing Pre-K participants to a comparison group of students who did not attend state-funded Pre-K. The primary objective of the project is to assess whether children who attended a Tennessee-funded Pre- K program perform better academically in the short and long term than a comparable group of peers who did not attend Tennessee's Pre- K program. The evaluation will also investigate whether various characteristics of Tennessee's Pre- K programs impact short- and long-term achievement among students who attended these programs.[4]

Given the scope of this evaluation, the study will be conducted over a period of three years. Ten cohorts of students will be tracked through the Fifth Grade in order to capture both short-term (Grades K-2) and long-term (Grades 3-5) academic achievement; this process will begin with the cohort of students who attended Pre-K in the 1998-1999 academic year. At the completion of the evaluation, which will span the school years 1999-2000 through 2008-2009, there will be a total of five cohorts that have the potential to have assessment data spanning from Kindergarten to Fifth Grade, and data for at least one grade level for the other five cohorts.[5] The results from analyses of these cohorts will be covered over the course of six project reports (two annual, three interim, and one final report).

The objective of this Second Interim Report is to build on the results presented in the First Interim Report and to provide the Office of Education Accountability (OEA) and the Tennessee General Assembly a report on the results of an analysis of student outcomes for students who participated in Tennessee's Pre-K program between 1998-1999 and 2005-2006. Three academic years are covered in this Second Interim Report: 2004-2005, 2005-2006, and 2006-2007. Table 1 summarizes the years and grade levels under study for this report (note that other interim reports in this evaluation include data for other years—see page 58 for a summary of additional reports planned as part of this evaluation).

---

[3] Throughout this report, the term "Pre-Kindergarten and its abbreviation "Pre-K" are used to specifically refer to Tennessee's state-funded Pre-Kindergarten program and not any other type of early childhood education program. The term "non-Pre-K" is used to refer to students who did not attend Tennessee's Pre-K program, although they may have participated in other early childhood education programs.

[4] The present report explores student-level characteristics as predictors of student outcomes, including race, gender, Free/Reduced Price Lunch status, and other characteristics. Analysis of program-level characteristics as predictors of student outcomes depends on the available of data. More discussion with the Office of the Comptroller and the Office of Early Learning is necessary to identify appropriate variables to include in a program-level analysis as well as to identify a data sources for these variables. If possible, such analysis will be included in a subsequent report.

[5] This does not imply that the analysis will include data for all students for all years, as many students are not assessed in Kindergarten, First, and Second grade.

Table 1. Grade Levels and School Years Covered in this Report

	Pre-K Year	2004-2005	2005-2006	2006-2007
Cohort 1	1998-99	5th		
Cohort 2	1999-00	4th	5th	
Cohort 3	2000-01	3rd	4th	5th
Cohort 4	2001-02	2nd	3rd	4th
Cohort 5	2002-03	1st	2nd	3rd
Cohort 6	2003-04	K	1st	2nd
Cohort 7	2004-05		K	1st
Cohort 8	2005-06			K

Statistical analyses explore the short- and long-term impact of Pre-K participation on student assessments in Kindergarten, First, Second, Third, Fourth, and Fifth Grades. This report offers a detailed overview of the research methodology used to assess the effectiveness of Pre-Kindergarten programs in Tennessee and also provides a descriptive overview of Pre-Kindergarten students and programs in Tennessee (see Appendix A).

Background and Objectives of the Present Study

According to Request for Proposal (RFP) number 307.14-004 issued by the State of Tennessee, Office of the Comptroller, the intent of this evaluation is "to assess the effectiveness of Tennessee's pre-kindergarten program on student achievement. The assessment shall include analysis of near term effects (Kindergarten through Second Grade) and long term effects (Third Grade through Fifth Grade). Effectiveness shall be assessed by analyzing data…gathered and maintained by local education agencies and/or the state department of education."

Preliminary studies investigating the performance of students who participated in Tennessee's Pre-K program suggest that Pre-K participants may perform better in the short term (defined as performance in Grades K-2) as well as in the long term (defined as performance in Grades 3-5) than other at-risk students who did not attend state-funded Pre-K.[6] The present study has been commissioned as an independent investigation and seeks to assess the progress of Pre-K students over time relative to a comparison group of students who have similar characteristics but did not attend Pre-K.

Research Design

For the purpose of this project, and as specified by RFP 308.14-004, "Pre-Kindergarten students" refers to students who attend state funded Pre-Kindergarten programs; specifically, either the pilot Pre-Kindergarten programs or lottery/general fund-funded Pre-Kindergarten programs. Also for the purpose of this project, as defined by the RFP, the non-Pre-K

[6] An overview of the TDOE analysis is available for download from the Internet at: http://www.tennessee.gov/education/prek/doc/prek_in_tn.pdf.

comparison groups consist of students who do/did not attend Pre-Kindergarten but whose characteristics otherwise match as nearly as practicable those of "Pre-Kindergarten students."

This evaluation, again as specified by the State of Tennessee, Office of the Comptroller, utilizes a quasi-experimental research design known as the **nonequivalent groups design**. This methodology, although not without limitations, permits a comparison of Pre-K participants to a comparable group of students who did *not* attend state-funded Pre-K. This particular type of analysis is deemed to involve "nonequivalent groups" to acknowledge the fact that it does not involve random assignment of students to groups at the time of enrollment in Pre-K.[7] However, it is important to note that this design does not preclude the possibility of obtaining comparable groups through random selection. Additionally, it allows for a longitudinal assessment of the progress of both Pre-K and non-Pre-K participants over time.

Parents elect for their children to participate in the Pre-K program in Tennessee, and program eligibility is determined by state policy such that all children meeting the state-determined eligibility requirements may be served.[8] Thus, randomization was not utilized in the present study in terms of assigning students to the Pre-K group. This is an important consideration in understanding and interpreting the results of the present study, and in distinguishing the present research methodology from experimental research methods.[9] Random assignment to a treatment or control group effectively equates the groups *before* an intervention is administered (for example, participation in a Pre-K program) and helps ensure that any resulting differences between the groups in later measurements are due to the intervention under study and not some other systematic difference between the treatment and control group. Experimental research methodology uses random assignment to create treatment and comparison groups—that is, the researchers conducting the study determine on a randomized basis which participants receive the treatment (the experimental group) and which do not (the control group). The experimental method is considered the most rigorous of research designs and enables researchers to address cause-and-effect relationships with the greatest degree of certainty.[10]

However, when implementing and evaluating complex educational programs, experimental methods are not always the most practical choice. First, fledgling programs often devote their resources to program implementation first and incorporate evaluation later. Thus, new programs are rarely designed with a rigorous experimental evaluation in place at the beginning. Further, researchers simply cannot control all the important variables which are likely to influence program outcomes, even with the best experimental design. Educational programs do not operate in a vacuum; even with a rigorous experimental design, researchers cannot be completely confident that any individual program independently produces specific results in terms of student achievement.[11] Thus, although utilizing random assignment is advantageous it does not in itself guarantee high internal validity—and may actually create a "false sense of

---

[7] Cook, T.D. & Campbell, D.T. (1979). Quasi-Experimentation: Design and Analysis for Field Settings. Rand McNally, Chicago, Illinois.

[8] See Appendix A for program overview including eligibility requirements.

[9] Campbell, D. T., & Stanley, J. C. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

[10] Trochim, William M. The Research Methods Knowledge Base, 2nd Edition. Internet WWW page, at URL: <http://www.socialresearchmethods.net/kb/> (version current as of October, 2006).

[11] Gribbons, B., & Herman, J. (1997). True and quasi-experimental designs. Washington, DC: ERIC Clearinghouse on Assessment and Evaluation. [ED421483]

security" in the research findings.[12] Experimental designs tend to be rare given the complexity and expense required to implement them effectively and because of logistical and ethical concerns—for example, is it ethical to deny a child access to an intervention like Pre-K?

Because of such limitations, other designs like the quasi-experimental design utilized in the present evaluation are often reasonable alternatives to address research questions of interest. Although quasi-experimental designs do not possess the same degree of scientific rigor as the experimental design, they are a practical and frequently utilized technique in applied social science.

In the present study, rigorous sampling techniques were used to select a comparison group from the many Tennessee schoolchildren who completed assessments in Grades K-5 but did not attend Pre-K, with the aim of constructing a valid comparison group that is matched as practicably as possible with the Pre-K group. Still, by the very nature of this research design, there is no way to ensure that the groups are, indeed, equivalent in all respects (thus the use of the term "nonequivalent groups"). There may be important differences between the Pre-K group and the non-Pre-K participants that simply cannot be captured retrospectively and accounted for in the data available for analysis in this report. Further, we can safely assume that there are important ways the non-Pre-K students may differ from the Pre-K participants. For example, a student may not have participated in Pre-K but may have participated in some other form of early childhood educational intervention. Unfortunately, the data available for analysis at present do not address participation in other early childhood programs and thus we cannot statistically control for the possibility that non-Pre-K participants did not receive *any* other form of intervention—we can only say for certain that they did not participate in Tennessee's Pre-K program. Random sampling, however, is the best technique to minimize the effects of such extraneous variables.

It is important to note that even if groups were constructed based on random assignment to the Pre-K and non-Pre-K groups, it would still be important to address whether non-Pre-K children participated in another, different early childhood education program. Ideally, at the time the groups were formed, information would be collected from both groups about their experiences. Because the present study is retrospective as opposed to prospective, there is a great deal of information about the comparison group that remains unknown. However, the goal of the present study was to describe the performance of Pre-K students on TCAP assessments relative to students who did not participate in Pre-K using data collected and maintained by TDOE—not to collect such additional data—although future prospective studies may be able to include such additional controls.

Finally, we acknowledge that this study also faces the limitation of utilizing a "post-test only" approach. That is, no baseline or pre-test data are available for either the Pre-K group or the non-Pre-K matched sample over the time period studied in this report. Given that randomization in selecting children to participate in the program is not feasible, there is clearly no possibility of statistically controlling for baseline differences for the non-Pre-K comparison group. Thus, we must make the assumption that the Pre-K and non-Pre-K groups "started out" at a similar point prior to the opportunity to participate in Pre-K. However, it is entirely possible given the nonrandom formation of the Pre-K group that the two groups may have initially differed had a pre-test been administered. From an evaluation standpoint, this makes any differences

[12] Gribbons & Herman (1997).

observed in later assessments difficult to interpret, and any such differences must be interpreted with caution.

Despite the limitations of the present design, this particular design offers some distinct advantages. First, because multiple measurements are available for the Pre-K and non-Pre-K groups, the resulting analyses afford a better sense of the patterns of variability *within* each group over time as well as *between* each group over time. Second, this design permits an exploration of ten years of existing data without the need to collect additional data on past program participants, a time-consuming and costly process. **The present study is not a means of conclusively determining whether participation in the Pre-K program causes an improvement in students' later performance on standardized assessments, and to construe it as such would be to misinterpret the goals and methodology applied here.** A prospective, experimental study would be better suited to permit such conclusions about the program. However, using existing data collected and maintained by TDOE, the present study uses the data at hand to provide the most accurate description possible of how Pre-K participants are doing in the short- and long-term based on the information available at the present time. Thus, the overarching goal of the present evaluation is to identify dominant trends in the overall pattern of results for Pre-K and non-Pre-K students and to determine if, overall, Pre-K students demonstrate any clear differences over time in their performance on these assessments relative to the non-Pre-K comparison group.

Methodology

For the present study, the Tennessee Department of Education (TDOE) provided the following datasets: student assessment data from 2004-2005 through 2006-2007, a file of Pre-K attendees spanning 1998-1999 through 2005-2006, and student demographic information from TDOE's Education Information System (EIS) for 2005-2006 and 2006-2007.[13] Great care was taken by TDOE and SRG to ensure student anonymity. No identifying information was provided along with student outcome data. To conduct the present study, these data sources were merged, and any irregularities or inconsistencies between the sources had to be addressed and reconciled. The Methodology section of this report details the nature of each data source, how SRG combined them to construct samples of Pre-K and non-Pre-K participants, and how this process impacted our analytic approach.

Data Sources

For the Second Interim Report, SRG drew from three data sources: 1) Pre-Kindergarten demographic data, 2) K-12 student assessment data, and 3) EIS student data from the 2005-2006 and 2006-2007 school years.

1. Pre-Kindergarten Demographic File

The Pre-Kindergarten (Pre-K) demographic file is a database maintained by the TDOE's Office of Early Learning. It was provided to SRG via the Director of Data Quality for the TDOE. The database spans eight academic years from 1998-1999 to 2005-2006. Starting

[13] Again, we remind the reader that other years of data are analyzed in other interim reports included in this evaluation. See page 58 for a summary.

with the 2006-2007 school year, information about Pre-K students is included in the Education Information System (for more information about the EIS, see the following section).

The Pre-K database contains information on the school (including county, system/local education agency (LEA), and school/provider name), program information (e.g., Pre-K funding source), and student demographic information (date of birth, gender, race, free/reduced price lunch (FRPL) status, special education status, whether English is the student's native language, and whether the school provided transportation). Although information is not available for all variables for all years in the Pre-K demographic file, the most important function of this data source is to identify students who participated in Tennessee's Pre-K Program beginning in 1998-1999 through 2005-2006.

To protect student confidentiality and to comply with federal regulations regarding student FRPL status, SRG did not obtain student names or Social Security Numbers. Social security numbers, however, were encrypted by TDOE so that the various data sources could be combined for the data analysis. This permitted SRG to link student assessment results with student demographic information and Pre-K participation data, but in a way that maintained student confidentiality.

It is important to note here that data management for the Pre-K demographic file has been an ongoing process. As more data become available—that is, as additional years of assessment and EIS data are incorporated into the analysis—we are able to cross-check more Pre-K students who had questionable records in earlier files and attempt to resolve inconsistencies. This requires us to exclude some students over the course of the evaluation but enables us to include others who had to be excluded from previous analyses. This will be discussed further in the Data Management section. Table 2 below provides the numbers of Pre-K students (at age 4) each year who have a valid record in the Pre-K demographic file at this point in the evaluation.

Table 2. Grade Levels and School Years Covered in this Report

Year	Number of Pre-K Students
1998-1999	318
1999-2000	311
2000-2001	1,102
2001-2002	2,195
2002-2003	2,631
2003-2004	2,405
2004-2005	2,345
2005-2006	7,129

2. Education Information System Data

The Education Information System (EIS) is a web-based data repository containing detailed student, teacher, school, and district level information. All schools input information in a

standardized format, and the EIS system is designed to catch data entry errors. EIS data are available beginning with the 2005-2006 school year. Although EIS includes data for prior school years, SRG was informed that these data are not complete and the state-assigned student ID number was only implemented in 2005-2006.

EIS files were provided to SRG by the TDOE Director of Data Quality. The data are in the form of spreadsheets that include demographic information, attendance records, disciplinary records, and special education records. EIS contains data for students in Kindergarten through Twelfth Grade, and for Pre-K students beginning in 2006-2007.[14]

3. K-12 Student Assessment Data

The third data source available for this evaluation is standardized assessment scores for students. These files were provided to us by the TDOE Assessment, Evaluation, and Research Division via the Department's Director of Data Quality. SRG requested and received scores for the 2004-2005 through 2006-2007 school years. SRG will obtain scores for future years as they become available and necessary for subsequent stages in the analysis. The files contain: 1) demographic characteristics of students (e.g., date of birth, gender, race) and 2) test scores in the following general subject areas: reading/language arts, mathematics, science, and social studies, along with composite scores by academic year. [15]

The TCAP (Tennessee Comprehensive Assessment Program) is the principal tool for assessing the performance of public school students in the State of Tennessee. The TCAP includes Tennessee-specific assessments which allow students, parents, and educators to interpret test scores as they relate to Tennessee's state curriculum standards.

For students in Grades K-2, the TCAP currently consists of Norm-Referenced Tests (NRT). Students in Grades 3-8 currently take Criterion-Referenced Tests (CRT). NRTs measure student performance relative to other test takers. Comparatively, CRTs measure performance according to specific standards, and test items are directly linked to specific performance indicators in the state curriculum.

The test for Kindergarteners includes Reading, Language Arts, and Mathematics. At First Grade, the test includes Reading, Language Arts, Mathematics, Science, Social Studies, Word Analysis, Vocabulary, and Math Computation. The Second Grade test includes all these subjects and also incorporates Spelling. Administering assessments in grades K-2 is a choice determined by school systems, and systems who elect to administer these assessments must incur the costs for these assessments themselves. The CRT assessments, are required for all students in Grades 3-8 and include four subject areas: Reading/Language Arts, Mathematics, Science, and Social Studies.[16] Tennessee students are assessed each spring.

---

[14] SRG did not obtain data for students in Grades 6-12 as they are not needed for the present evaluation.

[15] See page 26 for a list of all specific assessments administered in Grades K-5.

[16] Note: The scope of the present analysis is focused on student performance in grades K-5.

Comparability of NRTs and CRTs

Although both NRTs and CRTs are important and valuable in their use and application, there are some issues in terms of their comparability. For example, when CRTs are employed, each individual student's results are compared with a predetermined standard. The performance of other students who also took the test at the same time is not taken into consideration in evaluating the results. Student scores are typically reported in terms of the number of items correct, or the percentage correct. In contrast, for NRTs, each individual student is compared with other students who took the test, and the score reflects that student's performance relative to other students (not a predetermined criterion). Scores are typically reported in terms of a percentile or stanine, which indicates the student's position relative to a national sample of other test-takers in the same cohort.

Because there are significant conceptual and practical differences in the nature of the CRT and NRT assessments, longitudinal analyses across these measures are not feasible. For this reason, we will examine short-term (Grades K-2) and long-term (Grades 3-5) outcomes among Pre-K and non-Pre-K participants separately.

Assessments Administered in Grades K-5

The TCAP Achievement test is mandated for all students in Grades 3-8. The test is not mandated for Grades K-2, however. School systems may elect to test students in Grades K, 1 and/or 2, and their choice to test may vary from year to year.

TDOE provided SRG a spreadsheet summarizing the number of assessments administered in Grades K-2 by each Local Education Agency (LEA) each year between 1998 and 2007. LEAs may administer tests for one, two, or three of these grade levels in a given year, and they may change their decision to administer assessments each year. Thus, there is a great deal of variability in the number of schools administering assessments for students in Grades K-2 across this time period.

TDOE also provided the total average daily attendance (ADA) for all Tennessee school districts each year between 1998 and 2007. Using the information provided by TDOE, SRG estimated the percentage of students who were assessed in Grades K-2 each year, and these estimates are summarized in Table 3. In the years covered in this report, only about 6-10% of students are assessed in Kindergarten, 35-42% as assessed in First Grade, and 62-80% are assessed in Second Grade.

Table 3. Estimated Percentage of All Tennessee Students Assessed in Grades K-2 between 2004-2005 and 2006-2007

	Grade K	Grade 1	Grade 2
2004-2005	10%	42%	80%
2005-2006	9%	38%	77%
2006-2007	6%	35%	62%

It is important to note that although a small percentage of students are assessed in Kindergarten, this still amounts to a relatively large number of students overall. For example, the 6% of Kindergarten students with assessment scores in 2006-2007 translates to just

over 4,000 students. And although students who attended Pre-K represent a relatively small subgroup of students overall, given that the Pre-K program has experienced continuous growth since its inception, the numbers of Pre-K students assessed in Grades K-2 for school years 2004-2005 through 2006-2007, will be sufficient to conduct longitudinal analysis, which was not possible for the First Interim Report.

SRG next proceeded with the process of identifying Pre-K students, locating their assessment results, resolving any data discrepancies or inconsistencies in the data sources, and drawing a comparable sample of students who did not attend Pre-K.

Data Management

As was mentioned previously, SRG requested and received assessment data for the 2004-2005 through 2006-2007 school years. The data were provided in two files: one containing the scores for the Norm-Referenced Assessments (administered to students in Grades K-2), and the other containing the scores for the Criterion-Referenced Assessments (administered to students in Grades 3-5). In the original datasets that were provided by the TDOE, there were 156,465 cases in the NRT dataset and 419,140 in the CRT dataset. The two datasets were merged together into one dataset, and readied for analysis, which required several steps.

1. **Identify Pre-K Students in the Assessment Data**

The first step in the data management process was to identify which students in the assessment datasets attended Pre-K. To do so, the assessment datasets were merged together with the Pre-K demographic file and a variable was created that indicated whether or not the student had attended Pre-K. This allowed us to individually examine questionable records of Pre-K students throughout the data management phase. The subsequent steps detail the effort taken to prepare Pre-K and non-Pre-K students' assessment records for analysis.

2. **Identify and exclude assessment records with duplicate encrypted Social Security Numbers (ESSNs).**

The next step in preparing the data for analysis was to identify and exclude records with duplicate encrypted Social Security Numbers (ESSNs). Each year the assessment data contained a small number of cases with duplicate ESSNs, meaning that there were two (and in a very small number of instances, three) sets of scores for the same grade level and school year linked to the same ESSN. An examination of duplicate records found that in most cases, although the ESSN was the same, the demographic information (i.e., date of birth, gender, and/or race) was not, indicating that the assessment scores were for different students. For students with duplicate records who had attended Pre-K, each record was individually cross-checked with the demographic information linked to the ESSN with the Pre-K demographic file (when available) and EIS data (again, when available) to determine which record was incorrect. For Pre-K students whose demographic information was not reported in the Pre-K demographic file and did not have a record in the EIS in 2005-2006 or 2006-2007, SRG provided the records to the TDOE, who un-encrypted the records and attempted to determine which were correct. For non-Pre-K participants, because of the relatively large number of cases, it was not feasible to cross-check all duplicate records individually against the EIS. In these cases both records were excluded from analysis. It should be noted however, that for most grade/years, cases with duplicate ESSNs represented a very small proportion of all cases for that grade/year.

3. **Identify and flag records for students with assessments scores for more than one grade level in the same school year.**

The third step was to identify and flag records for students that had assessment scores for more than one grade level in the same school year. Although it is reasonable for a student to have scores at the same grade level for consecutive years (e.g., scores as a First Grader in both 2004-2005 and 2005-2006) as a result of retention, multiple sets of scores in the same school year at different grade levels is indicative of an error.[17] An examination of a number of these instances found that in each instance, the two sets of scores, although linked to the same ESSN, differed on demographic information. Again, efforts were made to retain as many valid Pre-K student records by individually cross-checking these students' records with EIS data and seeking the help of the TDOE for Pre-K students who did not have an EIS record. Because it was not feasible to individually check non-Pre-K records with multiple sets of scores in the same school year at different grade levels, these records were excluded from the analysis.

4. **Examine the consistency of demographic information across students' available set of scores.**

For students who were assessed more than once between 2004-2005 and 2006-2007, SRG researchers checked the consistency of date of birth, gender, and race across students' available sets of scores. The ideal circumstance, of course, is one in which the student's reported demographic information was the same across assessment records. This was indeed the case for the great majority of student records. Instances in which demographic information was different, however, needed to be investigated further.

There are three likely causes for discrepancies in demographic information. First, values for these characteristics may have been reported or entered incorrectly at one or more points in time. Incorrect values or data entry errors are particularly likely for date of birth, considering the value contains three pieces of information (day, month, and year). Second, the Social Security Number may have been reported or entered incorrectly for a student in one or more grade/years. Third, in the case of race, students may legitimately change the racial group with which they identify from one year to the next.[18]

All records for Pre-K students with discrepant values for date of birth, gender and/or race across sets of scores were examined individually. Their demographic information was cross-checked against the Pre-K demographic file (when available) and EIS data (again, when available). In some instances, it was clear that the discrepancy was an error; for example, instances where the date of birth was off by exactly one year. In instances where it was not obvious whether there was an error in the demographic information or in the ESSN, or the data were not available for cross-checking, SRG researchers submitted these records to the TDOE, who attempted to identify the source of the error and either correct the demographic information or the ESSN.

It was not feasible to individually examine the records of non-Pre-K students with discrepancies in demographic information across years of assessment scores. Although in most cases it is more likely that the discrepancy is an error in one record rather than the records belonging to different students, there is no way to be certain. Cross-checking the

---

[17] This was verified by the Senior Executive Director for the TDOE Office of Assessment, Evaluation and Research.

[18] This is also theoretically possible for gender, although extremely unlikely for the age group under study.

records with the EIS does not provide a definitive answer, either, given that some school systems pull students' demographic information for the assessment data directly from the EIS data. In general, SRG believes that a conservative approach is the best approach; that is, we feel it is better to exclude a case with valid information from the analysis than include one with invalid information. Thus, non-Pre-K students with discrepant demographic information across years of assessment scores were excluded from the analysis. There was one exception: students who had different values for race were retained, provided their values for gender and date of birth were consistent.

5. Examine the consistency of demographic information between the assessment data and EIS data.

An additional means of checking the validity of student records was to compare demographic information for students who had both assessment scores and a record in the EIS in 2005-2006 and/or 2006-2007.

Following the same approach outlined in step four, all records for Pre-K students with discrepant values for date of birth, gender and/or race in the assessment and EIS data were examined individually. Their demographic information was cross-checked against the Pre-K demographic file, when available. In instances where it was not obvious whether there was an error in the demographic information or in the ESSN, or the data were not available for cross-checking, SRG researchers submitted these records to the TDOE, who attempted to identify the source of the error and either correct the demographic information or the ESSN. The small number of non-Pre-K students with discrepant demographic information between assessment and EIS data were excluded from the analysis. As before there was one exception: students who had different values for race were retained, provided their values for gender and date of birth were consistent.

6. Additional data management needed for Pre-K student records

Having checked all students' records for consistency and questionable information, we next turned our attention to a data management issue specific to the Pre-K student records. Specifically, we identified the year that students were recorded as having attended Pre-K (specifically, at age 4) and examined whether that information was consistent with the grade(s) and year(s) in which they had assessment data, allowing for the possibility of grade retention, demotion, and skipping. In other words, students who had assessment scores as Kindergarteners in 2004-2005 would be expected to have attended Pre-K in 2003-2004. As another example, students who had assessment scores as First Graders in 2004-2005 would be expected to have attended Pre-K in 2002-2003, or 2001-2002 for those who repeated Kindergarten or were delayed in starting Kindergarten. In the majority of cases, the grade(s) and year(s) for which Pre-K students had assessment scores were consistent with the year in which they were identified as having attended Pre-K. There were, however, some discrepancies. First, there were a small number of cases where the year the student was recorded as having attended Pre-K was later than the grade and year in which they had assessment scores. For example, a student who supposedly attended Pre-K in 2002-2003 but had assessment scores as a First Grader in 2000-2001, a scenario which is not possible. There were also a small number of instances where a student had scores in a particular grade much earlier than expected given when they were recorded as having attended Pre-K; for example, a student with scores as a Fifth Grader in 2004-2005 but who supposedly attended Pre-K in 2001-2002. Just as we had for other types of data inconsistencies, we compared the demographic information reported in the Pre-K demographic file (when available) with the assessment data and EIS data (again, when available) and in many cases found that the information did not match. This indicates that

although the ESSN in the different data sources is the same, the information does not belong to the same student. These cases were checked by TDOE and resolved when possible. Cases that could not be corrected were excluded from analysis. It is important to note, however, that these instances were fairly rare.

Table 4 displays the final number of Pre-K students with assessment scores for each grade/year covered in this report. The table also includes the percentage of students assessed in a given grade/year based on the total number of four-year olds with valid records in the Pre-K demographic file the year students likely attended Pre-K. It is important to keep in mind that the percentages of students assessed in each grade year are estimates. They do not take into consideration grade retention, demotion, or skipping, any type of attrition (such as leaving the TN school system), or new students entering the TN school system.

The reader should also keep in mind that the numbers of Pre-K students in Table 4 is smaller than those presented in Table A1 in Appendix A, as Table A1 includes counts of all Pre-K students enrolled each year, including 3- and 4- year olds, and estimates for the 2007-2008 academic year. Table 4 reflects the number of *valid records* in the Pre-K demographic file and Pre-K assessment records available for analysis at the conclusion of the data management phase of this analysis. For example, students for whom a Social Security Number was not reported to TDOE were missing this information in the Pre-K demographic file; thus data from these students cannot be included in the analysis.

Table 4. Number of Pre-K Students in the Pre-K Demographic File and Number and Percentage of Pre-K Students Available for Analysis in Each Grade/Year Covered in this Report

Year & Number of Pre-K Participants in PKD File	Number and Percent of Pre-K Students Assessed in Each Grade/Year		
	2004-2005	**2005-2006**	**2006-2007**
1998-1999 N = 318	**Grade 5** 157 (49%)		
1999-2000 N = 311	**Grade 4** 291 (94%)	**Grade 5** 286 (92%)	
2000-2001 N = 1,102	**Grade 3** 634 (58%)	**Grade 4** 643 (58%)	**Grade 5** 640 (58%)
2001-2002 N = 2,195	**Grade 2** 1,380 (63%)	**Grade 3** 1,734 (79%)	**Grade 4** 1,752 (80%)
2002-2003 N = 2,631	**Grade 1** 1,000 (38%)	**Grade 2** 1,712 (65%)	**Grade 3** 2,333 (89%)
2003-2004 N = 2,405	**Grade K** 210 (9%)	**Grade 1** 1,027 (43%)	**Grade 2** 1,482 (62%)
2004-2005 N = 2,345		**Grade K** 181 (8%)	**Grade 1** 870 (37%)
2005-2006 N = 7,129			**Grade K** 615 (9%)

The number of Pre-K students with valid records who were assessed in a given grade and year varies widely, from 157 students in Grade 5 in 2004-2005 to 2,333 students in Grade 3 in 2006-2007. There are two main reasons for the range of group sizes beyond naturally occurring differences in the number of students who completed Pre-K each year.

First, as was mentioned previously, assessments in Grades K-2 are not mandated; only about 6-10% of Pre-K students were assessed in Kindergarten between 2004-2005 and 2006-2007, 37-43% were assessed in First Grade, and 62-65% were assessed in Second Grade. It is clear that some number of students changed LEAs, and some number of students may have entered Kindergarten late or repeated a grade, placing them in a different cohort from which they started. We also must consider whether Pre-K students may be more likely to attend school in LEAs that do not assess students in Grades K-2. For the most part, the proportion of Pre-K students who completed assessments in grades K-2 each year is very similar to the proportion of all Tennessee students who completed assessments

in those grade/years (see Tables 3 and 4), with the exception of fewer Pre-K students being assessed in Second Grade prior to 2006-2007.

A second factor impacting the number of Pre-K students in each grade/year, as was discussed previously, is that some students whose records indicated demographic discrepancies were excluded from analyses. Students were also excluded if they were found to have more than one set of scores in a particular school year at different grade levels. However, this resulted in the exclusion of a small number of cases.

It is important to note that even though a relatively small percentage of Pre-K students have assessment scores in Kindergarten, the number of students for whom valid assessment records are available is sufficient to be able to conduct statistical analysis. Also important to note is that all Pre-K participants with valid assessment records will be included in subsequent analyses.

After all due care was taken to identify and resolve questionable records in the data sources, the final step in preparing the data for analysis was to select a comparable sample of students from the assessment data for each grade/year who did not attend Tennessee's Pre-K program. The following section details the sampling strategy used to identify a comparable sample of non-Pre-K students.

Sampling Strategy

In order to evaluate the short- and long-term impact of Pre-K on student outcomes, Pre-K students must be compared to a similar group of students that did not attend Tennessee's Pre-K program.

Just as with the First Interim Report, we selected the matched non-Pre-K samples such that they mirror the Pre-K groups with regard to gender, race, and FRPL status. For the First Interim Report we also matched the two groups on school district. Because the numbers of Pre-K students in each grade level were significantly larger in the years covered in this Second Interim Report, it was possible to match the non-Pre-K and Pre-K students first at the school level and then at the district level in instances where a match was not possible at the school level but was possible at the district level. This modification to the sampling strategy offers a greater degree of assurance that the Pre-K and non-Pre-K students are similar in key ways aside from individual characteristics (e.g., gender, race, and FRPL status).

Because there is an opportunity for longitudinal analysis for this Second Interim Report (due to increased numbers of Pre-K students and the inclusion of more groups of students in Grades 3-5 with more assessment data available), a second modification was made to the sampling strategy. Rather than select a different non-Pre-K sample for each grade and year, the goal was to match students as early as possible in the years covered in the report (again, 2004-2005 through 2006-2007) and follow the matched groups when possible (e.g., Grades K-2 and 3-5). For example, students who were assessed in Kindergarten in 2004-2005 and had attended Pre-K were matched with a sample of non-Pre-K Kindergarten students who were also assessed in 2004-2005. Pre-K students who were assessed in Grade 1 in 2005-2006 but who did not have assessment scores for the previous year were matched with a group of non-Pre-K students who also did not have assessment scores the previous year. To follow this example out to conclusion, Pre-K students who were assessed in Grade 2 in 2006-2007 but who did not have assessment scores for the previous two years were matched to a group of non-Pre-K Grade 2 students who also did not have assessment scores the previous two years.

This strategy allows us to track matched students over multiple assessment points (again, when available). Following the same matched students for multiple assessments (i.e., over multiple years) will reduce the problems associated with unobserved heterogeneity, or how individuals may differ in some way that we cannot determine or predict. By following the same students over time there is less concern that differences in performance from one year to the next are due to the group of students in one year being somehow different from the students in the following years. This strategy also increases the comparability of the Pre-K and non-Pre-K groups. For one thing, Pre-K and non-Pre-K students will be comparable with regard to exposure to assessment tests, at least for the time period under study.

A third revision to the sampling strategy was to include Pre-K and non-Pre-K students who had received special education services at any time in the three years covered in this report. Given that receipt of special education services may be correlated with academic achievement and possibly Pre-K attendance as well, we determined that controlling for receipt of special education services in the analysis was a better approach than excluding special education students from the analysis as was done in the First Interim Report.

The sampling strategy for the non-Pre-K samples involved creating a distribution of the Pre-K group for each grade/year by district, then by school within district, then by FRPL status within each school, then by race and gender within each school. The goal was to create a sample of non-Pre-K students that resembled the Pre-K students as closely as possible in terms of their school district, school, FRPL status, race, and gender by finding an appropriate number of non-Pre-K students with the same demographic characteristics as each individual Pre-K student (i.e., precision matching). With regard to the sizes of the non-Pre-K samples, we used a variable ratio strategy such that if the Pre-K group for a given grade/year was less than 250, we took a 4:1 sample (i.e., four non-Pre-K matches for every Pre-K student), a 3:1 sample when the Pre-K group was between 250-499 students, a 2:1 sample for Pre-K groups between 500-999, and a 1:1 sample for Pre-K groups 1,000 and larger. We chose this approach rather than a fixed sample size strategy (used in the First Interim Report) for two main reasons. First, although having equal sample sizes in both groups (Pre-K and non Pre-K) is ideal, when a sample size is small (e.g., less than 250), as is the case for the Pre-K group for a couple grades and years under study, there is a greater likelihood of having insufficient data to evaluate the outcomes of interest accurately. Selecting a comparison group that was at times larger than the Pre-K group (especially when the Pre-K group size is small) ensured adequate information was available for evaluation of these outcomes. However, as the Pre-K group size increased, the need for additional comparison observations decreased; that is, there was sufficient information to evaluate the outcomes of interest accurately. Thus, as the Pre-K group size increased, the results could be based on equal initial sample sizes for comparison groups for the largest Pre-K groups. The second reason a variable ratio selection criterion was utilized was to maintain a comparison group that was never more than four times larger than Pre-K group, an important consideration given that the overall population of students who did not attend Pre-K is much larger than the population of students who did attend Pre-K. This ensured that the results were not, in a sense, dominated by the comparison group.

Table 5 on the following page provides the Pre-K group sizes and corresponding non-Pre-K sample sizes for each grade/year. The reader will notice the precipitous decrease in Pre-K group sizes across several grade/years, particularly as students move through Grades 3-5. Such a decrease indicates that the majority of students in a given grade/year were assessed in a preceding year, and thus had already been matched and removed from the remaining pool of unmatched Pre-K students. The drop in numbers of unmatched Pre-K students as they progress through grades and years indicates that sufficient numbers of Pre-K students are assessed in

multiple years, facilitating longitudinal analysis. Such analysis requires students to have been assessed multiple times.

Table 5 also provides the percentage of Pre-K students for whom there was the appropriate number of non-Pre-K matches each grade/year, based on the sampling ratio. For example, in 2004-2005 there are 157 Fifth Grade students with assessment scores who attended Pre-K, so we attempted to match each Pre-K student with four non-Pre-K students (totaling 628 students). The resulting non-Pre-K sample size of 625 means 99.5% of the Pre-K students were matched. For most grade/years, a very high percentage of Pre-K students were matched with the appropriate number of non-Pre-K students. Compared with Grades 1-5, it was more difficult to identify matches for Pre-K students in Kindergarten, which is not surprising given that the pool of non-Pre-K students is smallest for this grade level because few LEAs administer assessments in Kindergarten.

Table 5. Pre-K Group Sizes, Non-Pre-K Sample Sizes, and the Percentage of Pre-K Students Matched for Each Grade/Year

2004-2005	2005-2006	2006-2007
Grade 5 Pre-K = 157 Non-Pre-K = 625 (99.5%)		
Grade 4 Pre-K = 291 Non-Pre-K = 853 (97.7%)	**Grade 5** Pre-K = 7 Non-Pre-K = 25 (89.3%)	
Grade 3 Pre-K = 634 Non-Pre-K = 1,246 (98.6%)	**Grade 4** Pre-K = 39 Non-Pre-K = 144 (92.3%)	**Grade 5** Pre-K = 7 Non-Pre-K = 23 (82.1%)
Grade 2 Pre-K = 1,380 Non-Pre-K = 1,366 (99.0%)	**Grade 3** Pre-K = 461 Non-Pre-K = 1,327 (96.0%)	**Grade 4** Pre-K = 34 Non-Pre-K = 95 (69.9%)
Grade 1 Pre-K = 1,000 Non-Pre-K = 968 (96.8%)	**Grade 2** Pre-K = 841 Non-Pre-K = 1,630 (96.9%)	**Grade 3** Pre-K = 612 Non-Pre-K = 1,188 (97.1%)
Grade K Pre-K = 210 Non-Pre-K = 554 (87.9%)	**Grade 1** Pre-K = 805 Non-Pre-K = 1,523 (94.6%)	**Grade 2** Pre-K = 552 Non-Pre-K = 1,057 (95.7%)
	Grade K Pre-K = 172 Non-Pre-K = 465 (67.6%)	**Grade 1** Pre-K = 695 Non-Pre-K = 1,331 (95.8%)
		Grade K Pre-K = 610 Non-Pre-K = 948 (77.7%)

To review, for each Pre-K student, we attempted to identify at random non-Pre-K students (again, one to four, depending on Pre-K group size) of the same race, gender, and FRPL status

within the same school, or else at least within the same district. Also, when it was necessary to choose a non-Pre-K match from an alternate school within the same district, preference was given to selecting students from schools where there were other students who had attended Pre-K. Although it was not always possible to match Pre-K students to non-Pre-K students in their own school, matching Pre-K students in schools where there were other Pre-K students helped maintain the comparability of the Pre-K and non-Pre-K groups.

It should be noted that non-Pre-K samples were drawn from a three-category classification of race (White, Black, and Other) rather than the five category classification available in the assessment data (White, Black, Hispanic, American Indian/Native American, and Asian/Pacific Islander). The very low numbers of students in the latter three categories meant that quite often these students could not be matched. Yet, it is important to maintain the minority status of these students through the creation of the "Other" category. Even after collapsing the three categories to create an Other category, there were still too few cases to allow them to be properly analyzed. For purposes of analysis, then, we created two categories for race—white and non-white.

Once the non-Pre-K groups were selected, chi-square tests were conducted to compare the groups to make sure the proportions of subgroups were equivalent with regard to gender, race, and FRPL status. Without exception, all Pre-K and non-Pre-K groups were found to be statistically equivalent.

At this point, as many Pre-K students as possible had been identified in the assessment data, and any inaccuracies or irregularities were resolved and a comparable sample of non-Pre-K students was selected for each grade/year in the timeframe under investigation. The next step was to conduct the appropriate statistical analysis to determine whether there were meaningful differences, in aggregate, between the Pre-K and non-Pre-K groups.

Analytic Approach

Once the Pre-K students had been identified in the assessment data, and once a comparable sample of non-Pre-K students had been selected, the next step was to move to the analysis of the assessment results. All data reported in subsequent tables include only valid student records for Pre-K students and the sample of non-Pre-K students.

Variables Included in the Models and Characteristics of Students

The following section provides the distribution of students for all of the key predictor variables in the analysis, for all students overall and also for the Pre-K group (8,507 students) and non-Pre-K group (15,344 students).

1. **FRPL status** (FRPL or no FRPL)**.** Students' FRPL status was coded into one of two categories. A student was identified as receiving FRPL if he or she received FRPL *at least once* in the time period under study (2004-2005 through 2006-2007) and/or while in Pre-K, for Pre-K students. The rationale is the same as used in the First Interim Report; although a student's status as receiving FRPL may change from one year to the next, this does not necessary imply a considerable change in socioeconomic status. Thus, to be conservative, we included students in the FRPL group if they had received FRPL status at any time from 2004-2005 through 2006-2007 and/or while in Pre-K. This also resolved the problem of missing FRPL data for any one year.

   Table 6 summarizes students' FRPL status overall and in the Pre-K and non-Pre-K groups.

Table 6. Free/Reduced Price Lunch (FRPL) Status for Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
FRPL	78.1%	79.5%	77.2%
No FRPL	21.9%	20.5%	22.8%
Total	100.0% (23,851)	100.0% (8,507)	100.0% (15,344)

2. Race (white/non-white). See page 20 for a discussion of this variable. It should be noted that there is a small number of Pre-K students for whom data on race were missing or race could not be determined due to conflicting information in the data sources. However, this represents just 1.8% of all the Pre-K students. Table 7 summarizes the proportion of white and non-white students in the Pre-K and non-Pre-K groups.

Table 7. Race of Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
White	63.4%	62.6%	63.9%
Non-white	36.6%	37.4%	36.1%
Total	100.0% (23,694)	100.0% (n = 8,351)	100.0% (n = 15,343)

3. Gender (male or female). Table 8 summarizes the proportion of male and female students overall and in the Pre-K and non-Pre-K groups.

Table 8. Gender of Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
Male	51.4%	51.4%	51.3%
Female	48.6%	46.8%	48.7%
Total	100.0% (23,850)	100.0% (n = 8,506)	100.0% (n = 15,344)

4. Special education status (yes/ received special education or no/did not receive special education). Similar to the FRPL measure, special education students were identified as those who had received special education services at any time from 2004-2005 through 2006-2007 and/or while in Pre-K. In the assessment data, students were categorized by the numbers of hours in which they received special education services per week. Students who

received any special education services, regardless of how many hours a week, were identified as having received special education services. Table 9 summarizes the proportion of students receiving special education services overall and in the Pre-K and non-Pre-K groups.

Table 9. Special Education Services Received by Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
Yes	15.5%	18.1%	14.5%
No	84.2%	81.9%	85.5%
Total	100.0% (23,851)	100.0% (n = 8,507)	100.0% (n = 15,344)

5. Retention (retained/not retained). Students were deemed to have been retained if they had more than one year of assessment scores in the same grade level. It is important to keep in mind that the variable only captures those students who were retained between 2004-2005 and 2006-2007 (the time period under study in the present report). Further, because these students were identified from their assessment data and not all students are assessed (particularly in Grades K-2), the percentage of students retained will be lower than if we had data for all students. Table 10 summarizes the proportion of students retained overall in the Pre-K and non-Pre-K groups.[19]

Table 10. Retention for Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
Yes	2.0%	1.8%	2.2%
No	98.0%	98.2%	97.8%
Total	100.0% (23,851)	100.0% (n = 8,507)	100.0% (n = 15,344)

6. Native English speaker (yes/native English speaker or no/non-native English speaker). Native English speakers are defined as students whose primary or native language is English. This information was obtained from the EIS. Because we only have EIS data starting with the 2005-2006 school year, if a child was only assessed during the 2004-2005 school year, no information regarding his/her native language was available.

[19] Because retention was determined based on TDOE assessment data, these figures likely underestimate the number of students who were actually retained—which may be as high as 18% according to some estimates. Retention is a critical outcome, but given the difficulty in determining an accurate retention rate based on state assessment records, a more thorough analysis of retention (based on EIS records) is planned for subsequent reports.

Table 11. Native English Speaker Status for Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
Native English Speaker	85.7%	85.8%	85.7%
Non-Native English Speaker	14.3%	14.2%	14.3%
Total	100.0% (19,245)	100.0% (7,261)	100.0% (11,984)

7. Attendance (number of full-day absences). The attendance variable (days absent) is a continuous variable that ranges from 0 to 135. The original data contained values greater than 135 but they were few (only 61 students total, or 0.8% of students were reported to have missed more than 135 days in a given school year) and the numbers ranged much higher than possible (e.g., 300+ days absent in a single year). Because extreme values could not be used and it was a priority to include all possible cases in the Pre-K and non-Pre-K samples, attendance was truncated, or capped at 135 days absent. Because this information was obtained from the EIS, no attendance information is available for students assessed only in 2004-2005. Table 12 summarizes average student attendance overall and in the Pre-K and non-Pre-K groups.

Table 12. Mean Annual Attendance for Pre-K and Non-Pre-K Students

	Overall	Pre-K	Non-Pre-K
Mean (days)	8.5	8.3	8.6
Total (students)	22,475	8,303	14,172

Characteristics of the Assessments

As indicated previously, there are some differences in the number and type of assessments administered each year in Grades K-2 and 3-5. Table 13 summarizes the assessments and the grade levels in which they are administered.

Table 13. Summary of Assessments Administered in Grades K-5

	Kindergarten	First Grade	Second Grade	Third Grade	Fourth Grade	Fifth Grade
	Norm-Referenced Assessments			Criterion-Referenced Assessments		
Language Arts	X	X	X			
Math Computation		X	X			
Mathematics	X	X	X	X	X	X
Reading	X	X	X	X	X	X
Science		X	X	X	X	X
Social Studies		X	X	X	X	X
Spelling			X			
Vocabulary		X	X			
Word Analysis		X	X			

To have a better understanding of the number of students assessed at multiple time periods (and thus able to be included in longitudinal analyses), Table 14 presents the number of students in the Pre-K and non-Pre-K samples who were assessed at one, two, or three time points in Grades K-2. Table 15 presents the number of Pre-K and non-Pre-K students assessed at one, two, or three time points in Grades 3-5.

Table 14. Number of Pre-K and Non-Pre-K Students Assessed One, Two, or Three Years in Grades K-2

	Overall	Pre-K	Non-Pre-K
One year	11,316 (70.4%)	4,333 (69.2%)	6,983 (71.1%)
Two years	4,245 (26.4%)	1,772 (28.3%)	2,473 (25.2%)
Three years	522 (26.7%)	160 (2.6%)	362 (3.7%)
Total	16,083 (100.0%)	6,265 (100.0%)	9,818 (100.0%)

Table 15. Number of Pre-K and Non-Pre-K Students Assessed One, Two, or Three Years in Grades 3-5

	Overall	Pre-K	Non-Pre-K
One year	7,352 (51.3%)	2,694 (51.1%)	4,658 (51.4%)
Two years	5,323 (37.1%)	2,002 (37.9%)	3,321 (36.6%)
Three years	1,666 (11.6%)	581 (11.0%)	1,085 (12.0%)
Total	14,341 (100.0%)	5,277 (100.0%)	9,064 (100.0%)

Most Pre-K and non-Pre-K students were assessed at least once in Grades K-2. As discussed previously, only about 18-21% of all Tennessee students are assessed in Kindergarten and about half are assessed in First Grade, and this explains the relatively small percentage of students who are assessed at three time points between Kindergarten and Second Grade. The reader may note that more students were assessed at three time points in Grades 3-5, although there are still relatively few who were assessed at all three time points in the years covered in this report. As can be seen in Table 1 (see page 6), there is only one cohort of students who had the opportunity to progress from Grades 3-5 during 2004-2005 through 2006-2007 (i.e., those who were in Grade 3 in 2004-2005) and two cohorts of students with two possible years of assessment data in Grades 3-5 (those in Grade 4 in 2004-2005 and Grade 5 in 2005-2006, and those in Grade 3 in 2005-2006 and Grade 4 in 2006-2007). Thus, the number of students assessed in multiple years is largely determined by the timeframe covered in this report.

Modeling Strategy

The data were analyzed using *random effects models,* also referred to as *hierarchical linear models* or *multilevel models.* These models allow for "nesting" in the data. Simply put, "nesting" occurs when observations are organized in multiple units or levels. This is commonly seen in educational data, such that multiple students may be "nested" in a particular elementary school. If multiple students are assessed from a subset of schools within a school district, the schools are also "nested" within a district. In this example children are nested within schools and schools are nested within school district.

It is important to consider these relationships because students in one aggregate unit are often more alike than students across different units. Consequently, student assessment scores from a particular elementary school will likely be more similar to one another (i.e., correlated with one another) than scores from children attending different elementary schools. This can occur because, all else being equal, children "nested" within the same school have a more similar learning environment than children from different schools. The same is true at the district level.

Another form of nesting occurs when a student is assessed multiple times (e.g., in different grades) using the same measure. Here is it likely that two assessment scores from a single child will be more alike and thus more highly correlated than two scores from two different children. In other words, Child A's Kindergarten reading score will likely be more highly correlated with Child A's first grade reading score than it would be with Child B's First Grade reading score. This relationship arises from a common individual history (e.g., Child A's history) that influences the scores in a similar fashion over repeated measures. In cases like this where repeated observations are nested within a child's record and this relationship is incorporated into a random effects model, the subsequent model is often referred to as a "growth" or "trajectory" model.

In essence, the models used here cluster related observations into unique groups thereby controlling for these intergroup relationships--for example, multiple observations from a single child are treated as a single group, or children who attended the same school may be treated as a unique group. Given this, the variability in scores can be decomposed into within-group and between-group variability. By doing so, the models provide a more accurate representation of the data. Indeed, failing to account for "nesting" can lead to biased findings and thus a misunderstanding of the processes giving rise to the observed scores.

The mean (i.e., average score) and variability of an outcome (i.e., how scores vary around the mean) are of interest in the models presented here. When nesting is present in the data, a portion of the variability associated with a given outcome is due solely to the similarities in the source (child, school, etc.) of the scores and not necessarily due to the predictors of interest (Pre-K participation). Failing to account for this nesting can lead to biased results—specifically, finding no effect of Pre-K when there was indeed an effect, or vice versa. Therefore, all models used in this report examined the degree of nesting and accounted for this dependency when required. More specifically, the models in the current report examined the relationship between each outcome and the predictors outlined above; see also Table 16 below. In order to obtain accurate estimates of the relationship between each of these predictors and each outcome, the models tested and accounted for multiple sources of score variability. These sources of variability included individual variability, school variability, and school district variability.

Table 16 summarizes the overall proportion of each of the key variables for the Pre-K and non-Pre-K groups and provides a "snapshot" of the student characteristics in the grade levels studied in this report. It is important to note that the purpose of Table 16 is not to compare the distributions of child characteristics across grades. Although the distributions of these variables

may appear to change somewhat from year to year, they simply reflect that the same students are not assessed each year; that is, groups of students are somewhat different for each grade level.

Table 16. Distribution of Student Characteristics by Grade Level

	Grade K[1] (N = 682)	Grade 1[1,2] (N = 2,294)	Grade 2[1,2,3] (N = 4,584)	Grade 3[1] (N = 5,634)	Grade 4[1,2] (N = 4,100)	Grade 5[1,2,3] (N = 1,774)
Group						
Pre-K	27.0%	46.4%	49.0%	46.2%	43.4%	33.7%
Non-Pre-K	73.0%	53.6%	51.0%	53.8%	56.6%	66.3%
FRPL status						
FRPL	77.6%	82.5%	81.2%	79.9%	78.9%	77.1%
Non-FRPL	22.4%	17.5%	18.8%	20.1%	21.1%	22.9%
Race						
White	86.7%	70.6%	64.3%	62.6%	60.0%	60.1%
Non-white	13.3%	29.4%	35.7%	37.4%	40.0%	39.9%
Gender						
Male	46.8%	50.4%	51.4%	51.9%	51.5%	50.9%
Female	53.2%	49.6%	48.6%	48.1%	48.5%	49.1%
Special Education						
Yes	25.2%	20.3%	19.2%	17.6%	17.4%	16.5%
No	74.8%	79.7%	80.8%	80.8%	80.6%	83.5%
Retained						
Yes	2.3%	1.6%	1.7%	1.3%	1.3%	1.4%
No	97.7%	98.4%	98.3%	98.7%	98.7%	98.6%
English native						
Yes	97.5%	91.5%	88.0%	87.3%	83.9%	81.5%
No	2.5%	8.5%	12.0%	12.7%	16.1%	18.5%
Days absent						
Mean days	9.5	8.7	8.3	8.0	7.7	7.2

Note 1: A superscript "1" denotes grades used in growth models, a superscript "2" denotes grades used in "difference" models, and a superscript "3" denotes grades used in "single" models. Actual values may change slightly for a given outcome due to missing observations.

Note 2: The percentages reported reflect the characteristics of all Pre-K students with valid records, as well as the matched samples of non-Pre-K students with valid records.

The General Model

All models evaluated for this report include the child's FRPL history and participation in Tennessee state-funded Pre-K as predictors of academic achievement. As will be seen in the discussion of the results, the results focus on these two child-level characteristics as well as their interaction. The models were structured in this way for theoretical and practical reasons.

The effect of Pre-K participation is of primary interest in this evaluation, and thus that is the central focus of all the analyses. Further, because the program specifically targets children deemed “at-risk” and FRPL status is the only consistent variable available for analysis that serves as a proxy for “risk,” FRPL status was also considered a variable of interest. FRPL status was not studied as part of the First Interim Report; rather, we only included in the analyses students who had received FRPL at least once in the years covered in that report. Given the large number of Pre-K participants who receive FRPL, as well as the theoretical significance of this factor on student outcomes, it was included in the present analyses.

Further, exploratory analyses conducted in the First Interim Report did find some effects associated with students’ gender and race; although these variables are not of foremost interest in the present study, it was also a goal to address them in the Second Interim Report as well. Thus, all models examined here controlled for a child’s race and gender. In addition, the models employed in this report also include additional control variables: whether or not a child received special education within the observed grades, whether or not a child was retained within the observed grades, the average number of days a child was absent from class during the observed timeframe, and whether or not the child’s primary or native language is English. These control variables (and their theoretically or statistically relevant interactions) were included to ensure an accurate representation of the population under study and to ensure potentially mitigating effects were accounted for in the model to control for any potential bias.

Depending on the number of grades a particular outcome was assessed, one of three models was used. These models include a *single time point model*, a *difference model*, and a *growth model*. Each of these models is described briefly below.

Single Time Point Models

When an outcome (i.e., assessment score) was observed in only one grade (for example, Spelling in Grade 2), the relationships between an outcome and predictors were examined using all children with a viable score on the outcome of interest who had no missing values for the predictors of interest and did not have multiple observations within a given grade; see Table 16 for the model sample sizes. The models also controlled for possible nesting within school and school district. The single time point models used in the present analyses are often referred to as random effects analyses of covariance, or ANCOVAs.

Difference Models (Two Time Points)

When an outcome was observed in two grades (for example, norm-referenced scores in Social Studies in Grades 1 and 2), the relationships between an outcome and the predictors were examined over the two grades using all children with a viable score on the outcome of interest in at least one grade. Children who had no missing values for the predictors of interest, did not have multiple observations within a given grade, and did not switch schools between assessments where included in the analyses; see Table 16 for the model sample sizes. These models, also called ANCOVAs, controlled for possible nesting within individual. While possible nesting within school and school district was examined, there was no evidence supporting the addition of these sources of variability to the model. That is, there was no statistical basis to include school and school district (LEA) as predictors in the model.

Growth Models (Three Time Points)

When an outcome was observed over three grades (for example, norm-referenced scores in Reading or criterion-referenced scores in Mathematics), the relationships between an outcome and predictors were examined over all three grades using all children with a viable score on the outcome of interest in at least one grade. Children who had no missing values for the predictors

of interest, did not have multiple observations within a given grade, and did not switch schools between assessments were included in the analyses; see Table 16 for the model sample sizes. These models controlled for possible nesting within individual within and over grade level. While possible nesting within school and school district was examined, there was no evidence supporting the addition of these sources of variability to the model. This type of model is often called a "growth model."

Comparing the Different Models

The difference between the three derivations of our general model outlined above lies in how each derivation incorporates (if it incorporates) information about the child's grade. The *single time point* model excludes child's grade from the model because the outcomes of interest were only measured in Second Grade. Thus, all scores for all children are from the same grade. The *difference model* did include grade as a predictor in the model. Given the outcomes of interest were only assessed over two grades (First and Second Grades), grade provided a way to measure the average change from First to Second Grade. Importantly, the inclusion of grade in the model also allowed for our predictors of interest and control variables to have differential effects over grade. For example, the inclusion of the grade by Pre-K interaction allowed for us to examine not just if Pre-K children differ from non Pre-K children within a given grade, but to also examine how their scores increase between grades. The *growth model* outlined above treats grade in a very similar fashion except the outcomes were now measured over three grades as opposed to two. Thus, grade here allows us to measure the average change between any two grades. Appendix B addresses the more technical aspects of the general model and how each of the three specific model types are related.

A standard level of α = .05 was used to determine statistical significance. This means that "statistically significant" differences indicate differences between the Pre-K and non-Pre-K groups that are highly unlikely to be observed simply due to chance. Adjustments were made to correct for multiple comparisons using the Benjamini-Hochberg (1995) method.[20] This adjustment was necessary because of the number of multiple comparisons conducted in the present analyses. Such an adjustment decreases the "false discovery rate," or the likelihood that a difference between groups will appear to be statistically significant, but is in fact a result of increased error. The Benjamini-Hochberg (B-H) procedure is a widely-accepted technique to control for false discovery and has been adopted for use in reporting results from the National Assessment of Educational Progress (NAEP),

Results

Organization of the Results

The results for this report are organized by the type of assessment (norm-referenced and criterion-referenced), and within each type of assessment by the number of time points studied (three, two, or one). For purposes of the present report, and as defined by the State of Tennessee, Office of the Comptroller, the results of Norm-Referenced Assessments administered in Grades K-2 reflect the short-term effects of Pre-K participation, and the results of Criterion-Referenced Assessments administered in Grades 3-5 reflect the long-term effects of Pre-K participation. For each assessment, model-implied adjusted mean scores are presented

[20] Benjamini, Y. & Hochberg, Y. (1995) "Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society*, Series B (Methodological), Vol. 57, No. 1, pp. 289-300.

for Pre-K and non-Pre-K students. These mean scores are adjusted for the variables included in the model, meaning that these scores control for gender, race, special education, retention, and number of full-day absences.

Within the analysis of norm-referenced tests (Grades K-2), the results are organized such that the results of growth curve models (spanning three time points) are discussed first, followed by difference score models (spanning two time points), and concluding with tests administered at only one point in time. The analyses conducted for Criterion-Referenced Assessments all include three time points (Grades 3-5), as assessments in Mathematics, Reading, Science, and Social Studies are administered to students in all three grades.

Organization of the Tables

Tables are organized such that overall differences between the Pre-K and non-Pre-K group are presented first (leftmost columns). Differences between Pre-K and non-Pre-K students who received FRPL are presented in the next two columns, and differences between Pre-K and non-Pre-K students who did not receive FRPL are presented last (rightmost columns). It is important for the reader to note that the primary comparisons of interest in this report are comparisons of the Pre-K and non-Pre-K groups, and so only statistically significant differences between the Pre-K and non-Pre-K groups are denoted in the tables and discussed in the results. Further, the reader should take caution that unless a difference between the Pre-K and non-Pre-K group is identified as "statistically significant," any differences between the Pre-K and non-Pre-K groups are likely (and expected to be) due to measurement error (e.g., inaccuracies in the data) and to natural group variation.

Although FRPL was included in these models to explore whether there were effects associated with Pre-K participation among FRPL and non-FRPL students and whether the results show similar patterns, we would like to reiterate that exploring outcomes associated with students' FRPL status is not a primary objective of this report. That is, our discussion of the results focuses on effects associated with Pre-K *within* categories of students' FRPL status. For example, any differences between the Pre-K and non-Pre-K groups are discussed and interpreted separately for students who received FRPL and for students who did not receive FRPL. This decision was made due to the important conceptual differences between the FRPL and non-FRPL groups. However, we acknowledge the importance of addressing FRPL status for practical and theoretical reasons, and so the effects of all other relevant control variables including students' FRPL status (i.e., whether a student did or did not receive FRPL in the time period under study) were included in order to help interpret the effects of Pre-K participation within the groups.

Norm-Referenced Assessments (Grades K-2)

Three assessments were administered to students in Grades K-2 at all three grade levels: Reading, Language Arts, and Mathematics. Scores were analyzed for students who completed these assessments in Kindergarten, First Grade, and Second Grade. The results of the growth curve modeling for all three assessments showed the same basic pattern of results. Beginning with Kindergarten, for Reading, Language Arts, and Mathematics, students' scores reflected a significant difference such that students who attended Pre-K (both those receiving FRPL and those who did not) tended to score higher than students who did not attend Pre-K. There was also a significant difference associated with students FRPL status. More specifically, among students receiving FRPL, Pre-K participants scored higher than their non-Pre-K peers. Among students not receiving FRPL, Pre-K participation was also associated with higher scores on these assessments. Thus, for Reading, Language Arts, and Mathematics, students who

participated in Pre-K tended to score higher, on average, in Kindergarten assessments (see Table 17).

The reader will note that Pre-K students who received FRPL scored similarly to students who did not attend Pre-K and who did not receive FRPL. This was the case for all three assessments (Reading, Language Arts, and Mathematics). In the models analyzed here, these differences were *not* statistically significant. While this does suggest that Pre-K participation does help to "close the gap" between students who are "at risk" due to socioeconomic status (i.e., the FRPL student group) and students "not at risk" as defined in this way, we would like to remind the reader that such an interpretation must be made with caution, as there are many ways to define "risk." Although we can say with some confidence that students who did not receive FRPL in the time period under study are not considered "at risk" by this definition, it does not mean that students in the non-FRPL group do not also possess some other characteristics that might be considered to place them "at risk." This particularly may be the case for Pre-K students, as priority is given to students who are deemed "at risk," in some way, even if they do not receive FRPL. So, to reiterate, this is a comparison that we strongly suggest should be interpreted with caution. Further, it also bears repeating that the groups were constructed based on whether or not a student participated in Pre-K, and so comparisons between the Pre-K and non-Pre-K groups within each subgroup (i.e., FRPL or no FRPL) are the most valid.

Table 17. Model-Implied Adjusted Mean Scores for Pre-K and Non-Pre-K Students—Kindergarten

	Group					
Assessment	**Pre-K**	**Non-Pre-K**	**Pre-K FRPL**	**Non-Pre-K FRPL**	**Pre-K No FRPL**	**Non-Pre-K No FRPL**
Kindergarten Reading	**567.01***	**557.64***	**560.69****	**553.85****	**573.32*****	**561.42*****
Kindergarten Language Arts	**571.93***	**560.00***	**564.17****	**554.25****	**579.69*****	**565.74*****
Kindergarten Mathematics	**529.51***	**516.29***	**519.31****	**512.00****	**539.71*****	**520.58*****

Note: Models based on a minimum sample size of 4,930 children.

* Denotes a statistically significant difference between Pre-K and non-Pre-K students, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

** Denotes a statistically significant difference between Pre-K and non-Pre-K students who received FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold)..

*** Denotes a statistically significant difference between Pre-K and non-Pre-K students who did not receive FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

As students' scores were modeled over the three time points for which data were available, a different pattern emerged for all three assessments in Second Grade. Although the Pre-K and non-Pre-K students differed in Kindergarten, they tended to converge into two groups by

Second Grade. In Second Grade, significant differences were no longer found as a function of Pre-K participation (see Table 18).

Table 18. Model-Implied Adjusted Mean Scores for Pre-K and Non-Pre-K Students—Second Grade

	Group					
Assessment	**Pre-K**	**Non-Pre-K**	**Pre-K FRPL**	**Non-Pre-K FRPL**	**Pre-K No FRPL**	**Non-Pre-K No FRPL**
Second Grade Reading	617.93	620.73	612.66	613.80	623.20	627.66
Second Grade Language Arts	626.05	628.88	620.72	621.35	631.38	636.41
Second Grade Mathematics	572.27	573.71	567.64	566.86	576.90	580.56

Note: Models based on a minimum sample size of 4,930 children.

* Denotes a statistically significant difference between Pre-K and non-Pre-K students, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

** Denotes a statistically significant difference between Pre-K and non-Pre-K students who received FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

*** Denotes a statistically significant difference between Pre-K and non-Pre-K students who did not receive FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

Change over time in students' average performance in Reading, Language Arts, and Mathematics in Grades K-2 is summarized in the following figures. Figures 1-3 show the results for the Pre-K and non-Pre-K groups overall, and Figures 4-6 show the results for Pre-K and non-Pre-K students within categories of FRPL status. It is important to note that the figures represent the **model-implied means**—the predicted means at each time of assessment. Given the number of time points and the nature of models employed in the current report, all model-implied means follow a linearly increasing or decreasing trend (i.e., a straight line).

Figure 1. Model-Implied Adjusted Mean Scores for Reading in Grades K-2 for Pre-K and Non-Pre-K Students

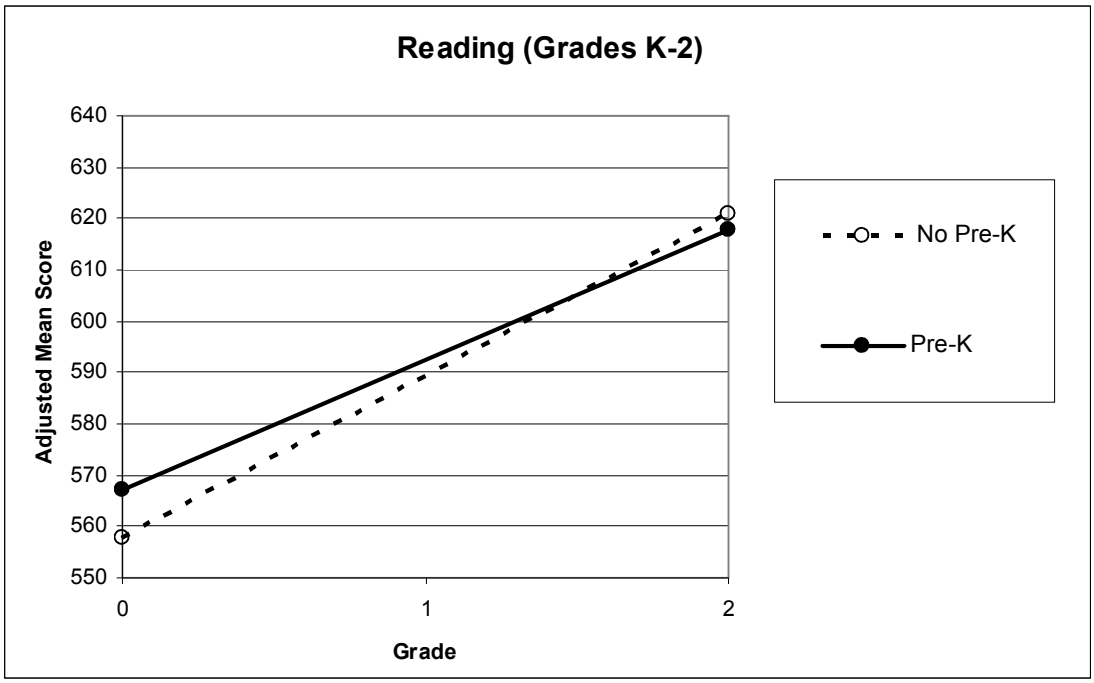


Figure 2. Model-Implied Adjusted Mean Scores for Language Arts in Grades K-2 for Pre-K and Non-Pre-K Students

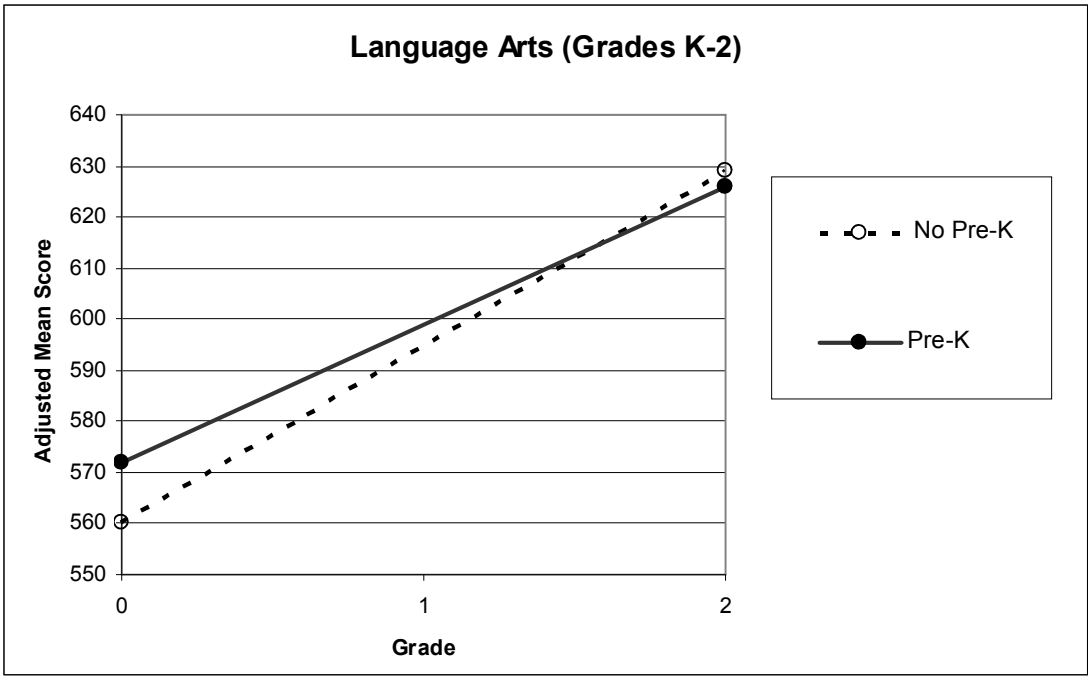


Figure 3. Model-Implied Adjusted Mean Scores for Mathematics in Grades K-2 for Pre-K and Non-Pre-K Students

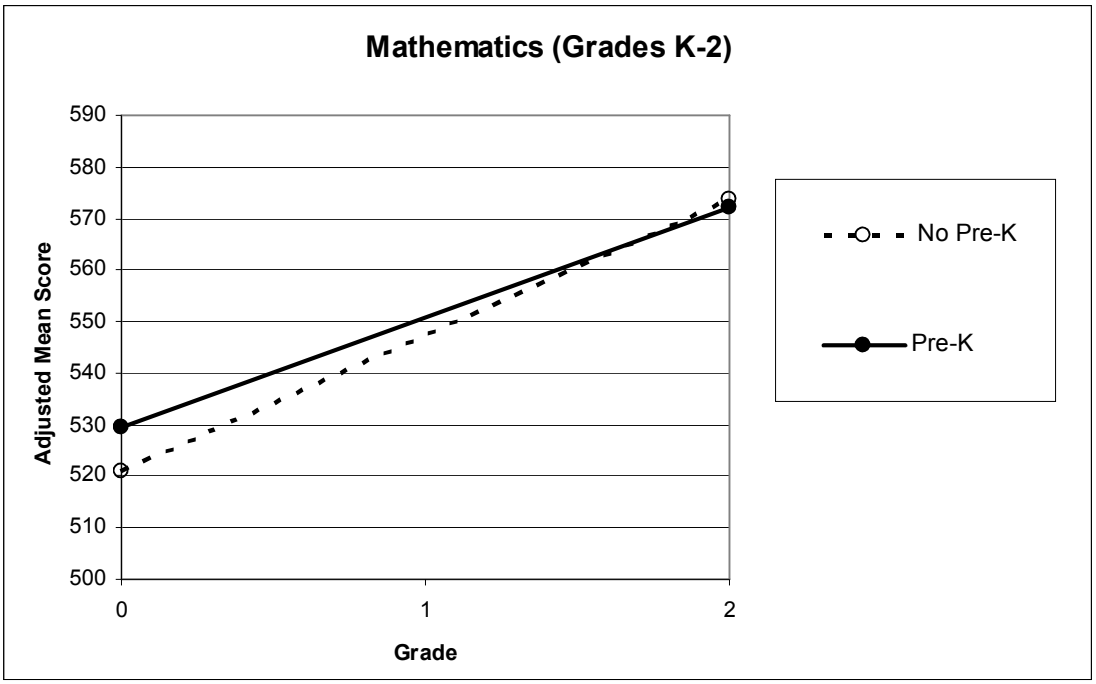


Figure 4. Model-Implied Adjusted Mean Scores for Reading in Grades K-2 for Pre-K and Non-Pre-K Students by Student FRPL Status

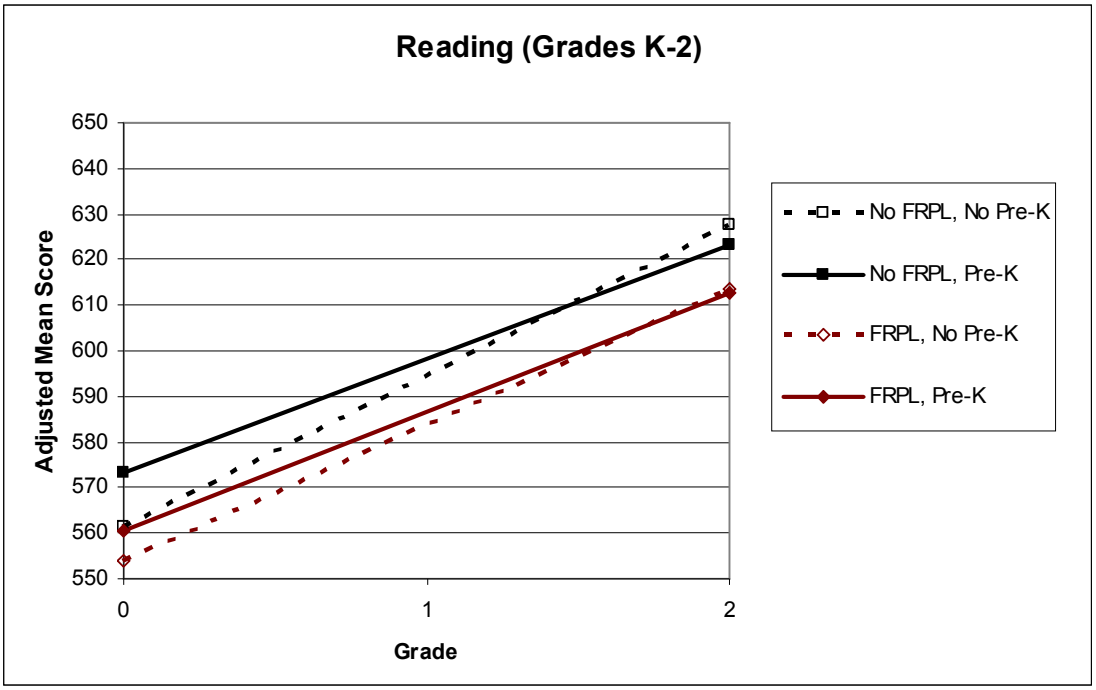


Figure 5. Model-Implied Adjusted Mean Scores for Language Arts in Grades K-2 for Pre-K and Non-Pre-K Students by Student FRPL Status

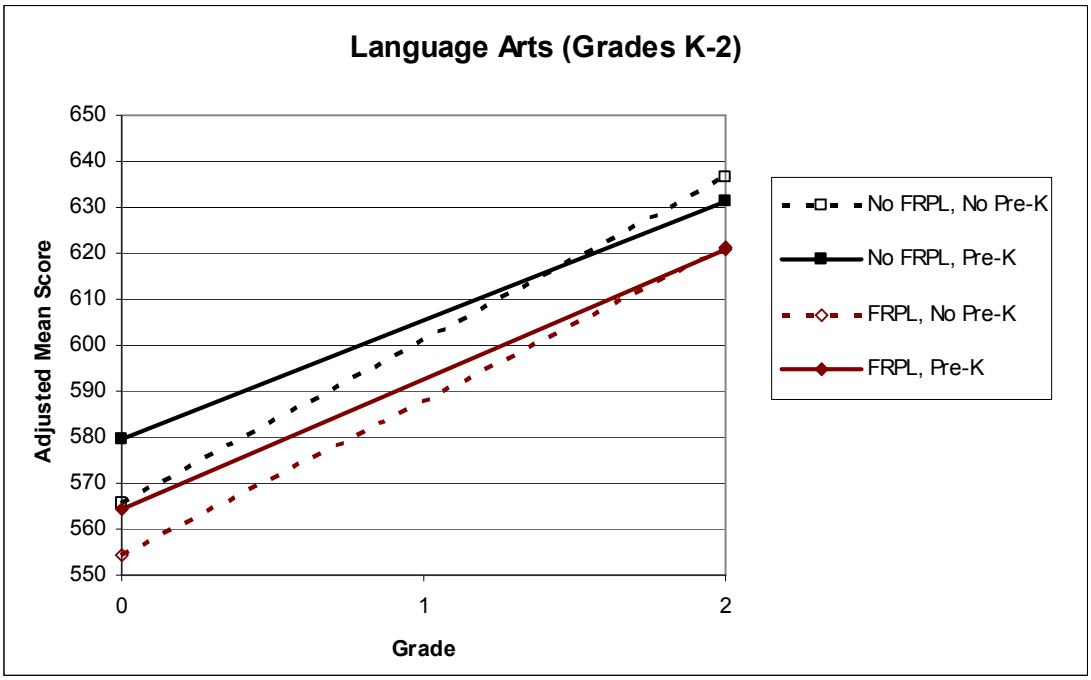
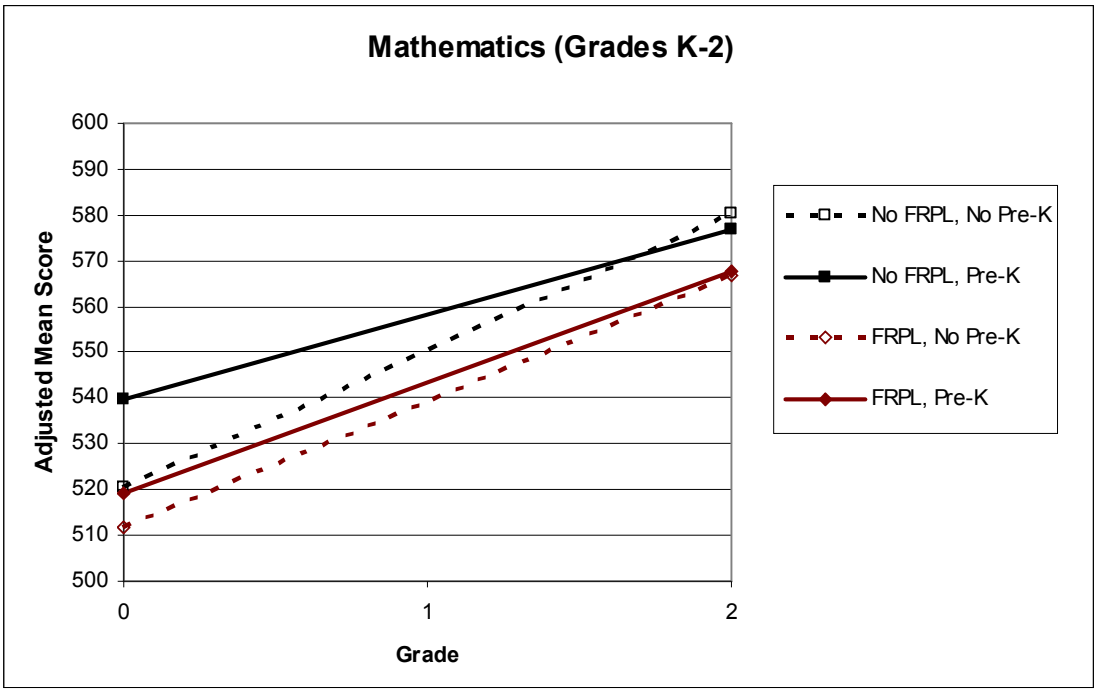


Figure 6. Model-Implied Adjusted Mean Scores for Mathematics in Grades K-2 for Pre-K and Non-Pre-K Students by Student FRPL Status



In addition to Reading, Language Arts, and Mathematics, students in the First Grade also complete Norm-Referenced Assessments in Vocabulary, Word Analysis, Math Computation, Social Studies, and Science. Because assessments are administered in both the First Grade and the Second Grade, it is possible to examine student performance over two years in these areas using difference score modeling.

The same pattern emerged across all assessments in both First Grade and Second Grade. In no case was Pre-K participation a significant predictor of student performance. Among students who did not receive FRPL, there were no statistically significant differences in First Grade scores for any of these assessments as a function of Pre-K participation, nor was this the case among students who did receive FRPL.

These results are summarized in Table 19. Figures 7-11 show the results for the Pre-K and non-Pre-K Groups in Vocabulary, Word Analysis, Math Computation, Social Studies, and Science assessments in First and Second Grade. Figures 12-16 show the results for these assessments including student FRPL status.

Table 19. Model-Implied Adjusted Mean Scores for Pre-K and Non-Pre-K Students—First and Second Grades

	Group					
Assessment	**Pre-K**	**Non-Pre-K**	**Pre-K FRPL**	**Non-Pre-K FRPL**	**Pre-K No FRPL**	**Non-Pre-K No FRPL**
First Grade Vocabulary	570.55	567.29	561.47	561.42	579.62	573.15
First Grade Word Analysis	597.85	594.88	591.34	588.65	604.36	601.11
First Grade Math Computation	505.70	504.67	503.44	501.31	507.96	508.03
First Grade Social Studies	596.67	594.72	591.68	591.08	601.65	598.35
First Grade Science	578.82	577.10	570.38	571.47	587.25	582.72
Second Grade Vocabulary	600.97	601.72	595.51	594.95	606.42	608.49
Second Grade Word Analysis	624.93	625.47	618.72	618.08	631.13	632.86
Second Grade Math Computation	549.46	550.20	545.04	544.66	553.88	555.73
Second Grade Social Studies	612.76	614.22	606.02	605.07	619.49	623.36
Second Grade Science	592.03	592.55	586.47	584.33	597.59	600.77

Note: Models based on a minimum sample size of 4,651 children.

* Denotes a statistically significant difference between Pre-K and non-Pre-K students, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

** Denotes a statistically significant difference between Pre-K and non-Pre-K students who received FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

*** Denotes a statistically significant difference between Pre-K and non-Pre-K students who did not receive FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

Figure 7. Model-Implied Adjusted Mean Scores for Vocabulary in Grades 1-2 For Pre-K and Non-Pre-K Students

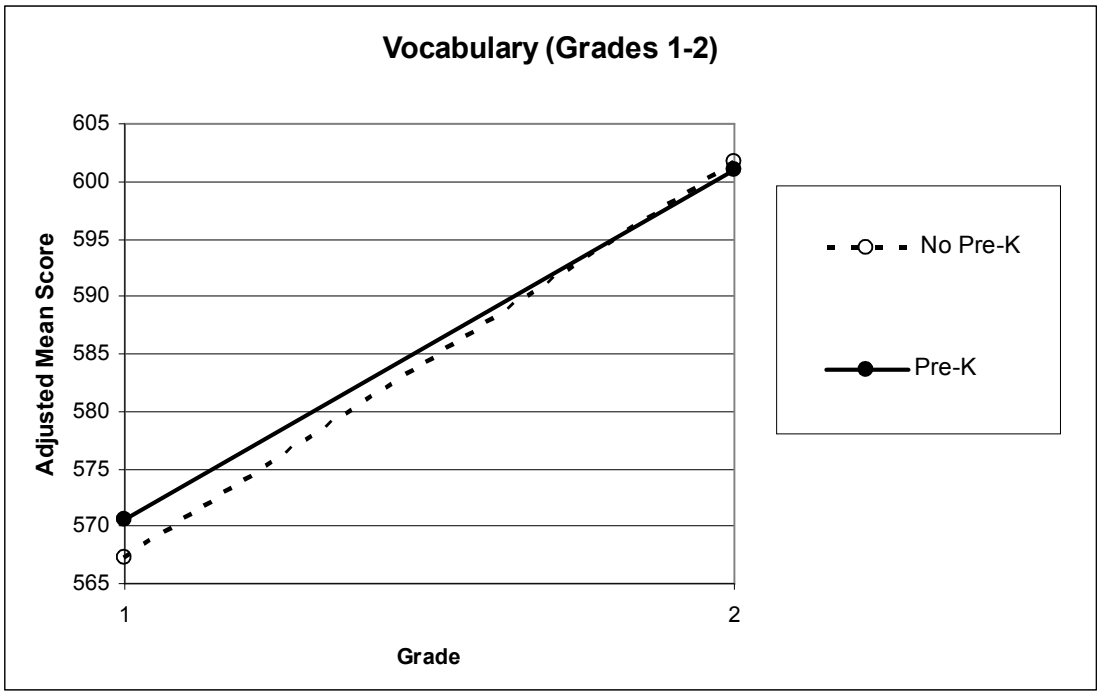


Figure 8. Model-Implied Adjusted Mean Scores for Word Analysis in Grades 1-2 for Pre-K and Non-Pre-K Students

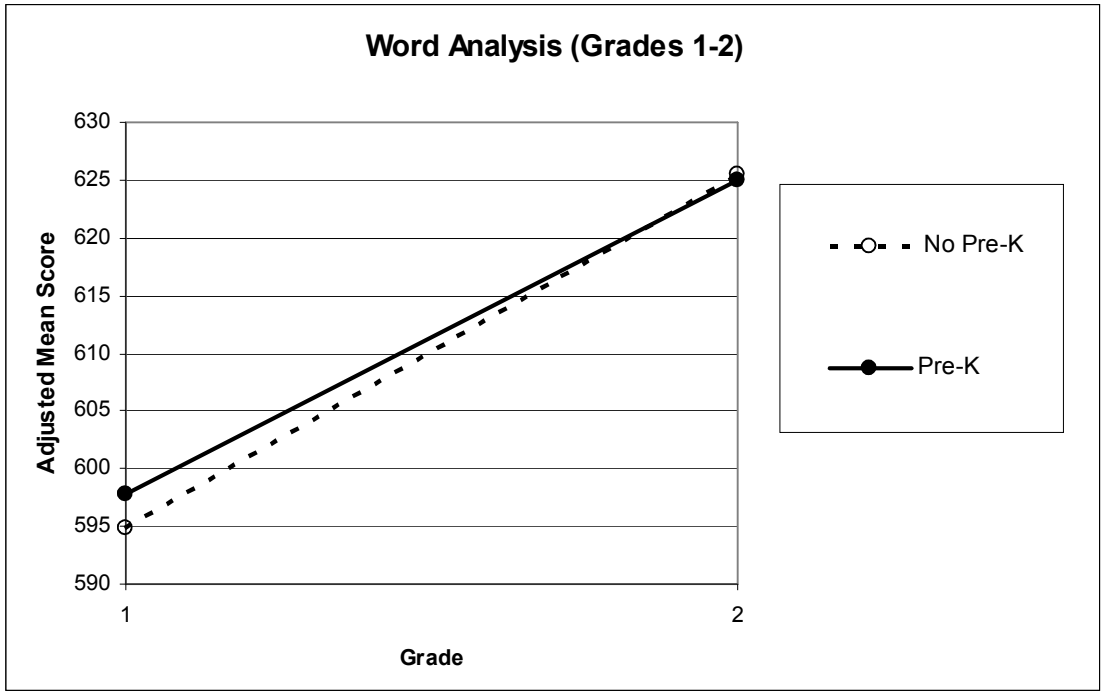


Figure 9. Model-Implied Adjusted Mean Scores for Math Computation in Grades 1-2 for Pre-K and Non-Pre-K Students

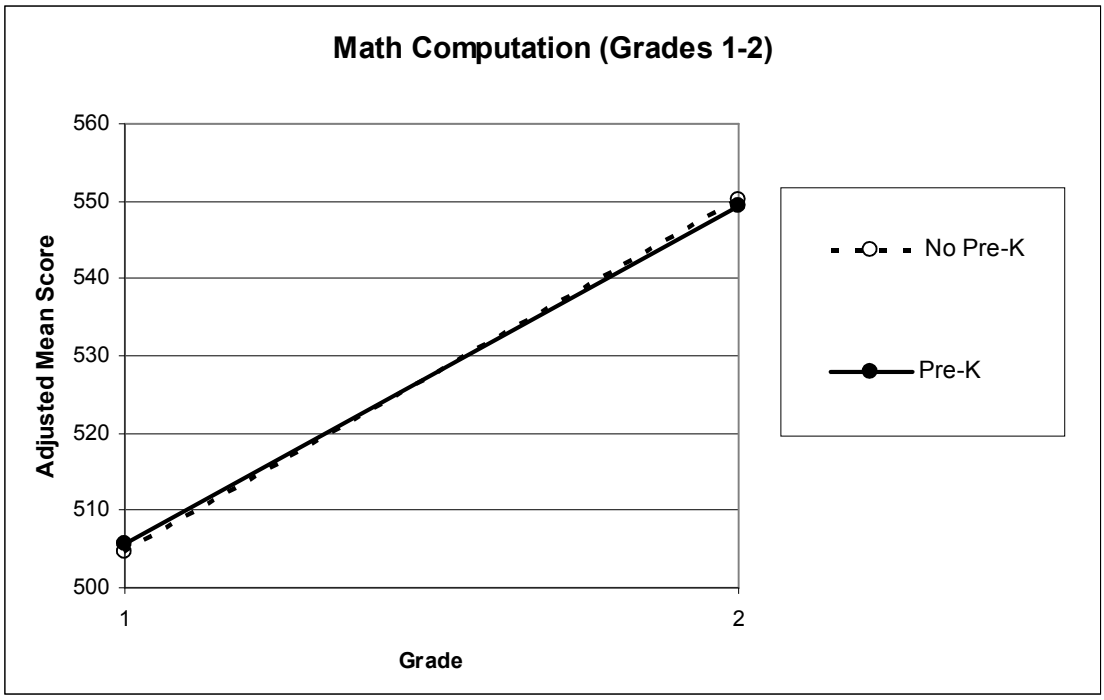


Figure 10. Model-Implied Adjusted Mean Scores for Social Studies in Grades 1-2 for Pre-K and Non-Pre-K Students

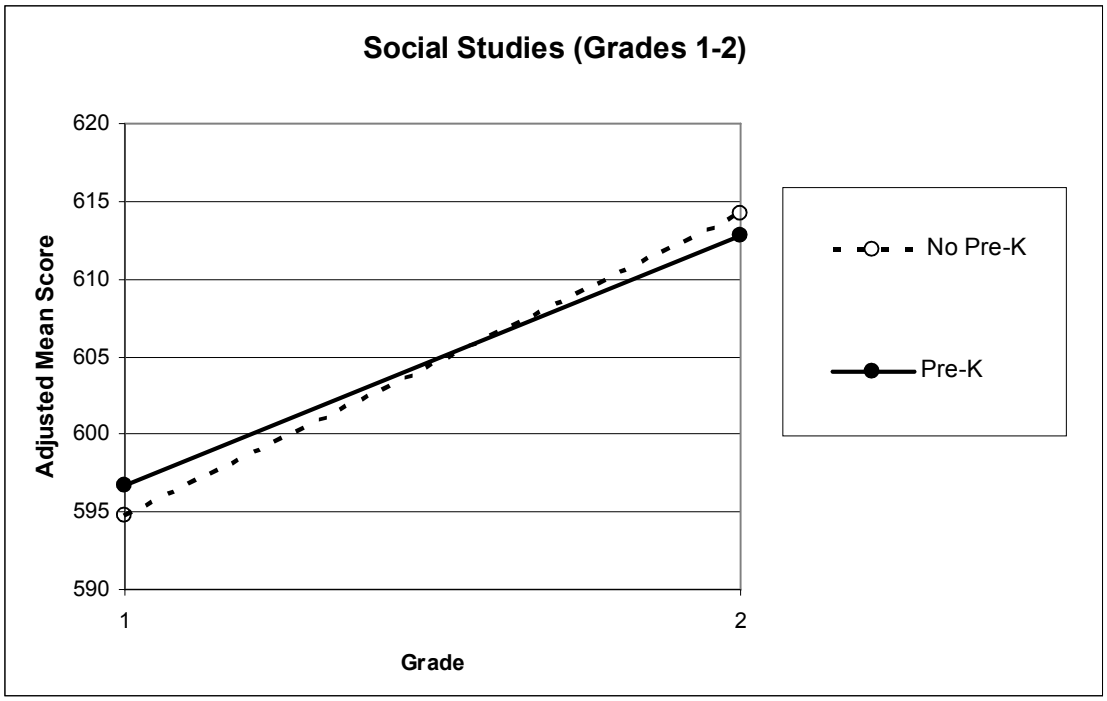


Figure 11. Model-Implied Adjusted Mean Scores for Science in Grades 1-2 for Pre-K and Non-Pre-K Students

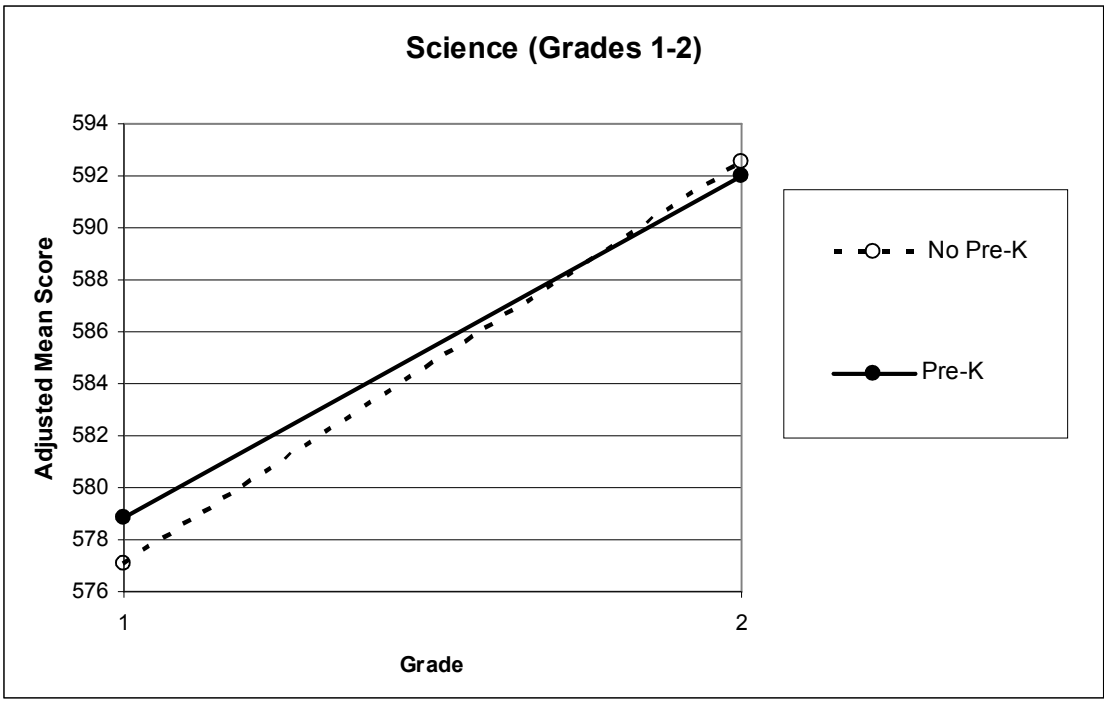


Figure 12. Model-Implied Adjusted Mean Scores for Vocabulary in Grades 1-2 For Pre-K and Non-Pre-K Students by Student FRPL Status

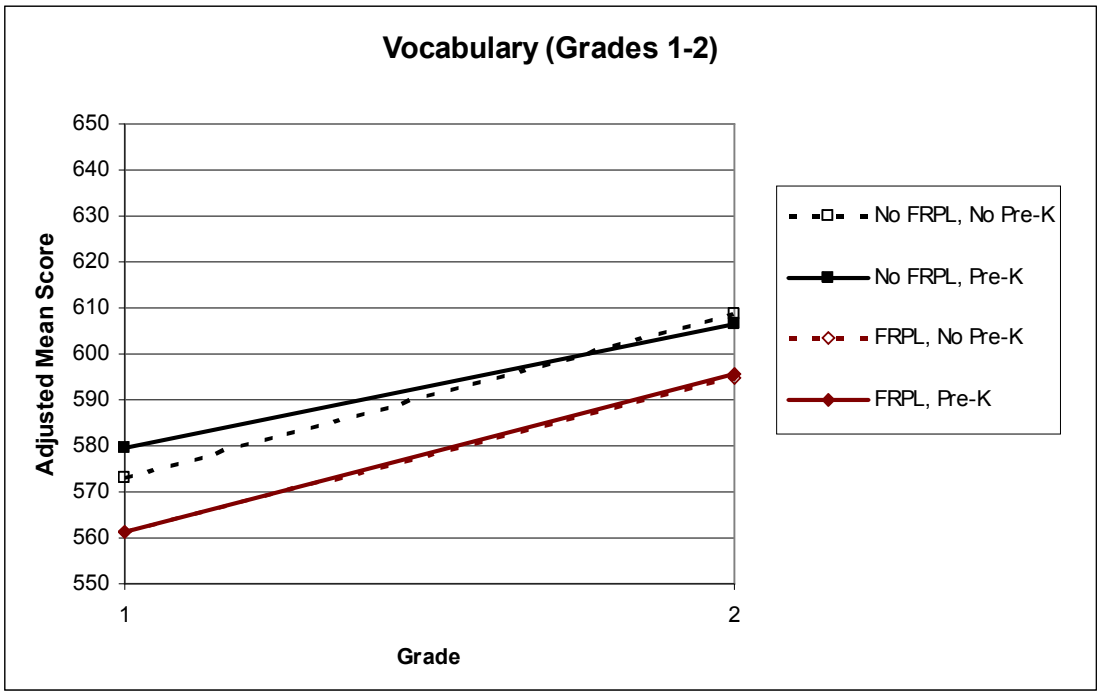


Figure 13. Model-Implied Adjusted Mean Scores for Word Analysis in Grades 1-2 for Pre-K and Non-Pre-K Students by Student FRPL Status

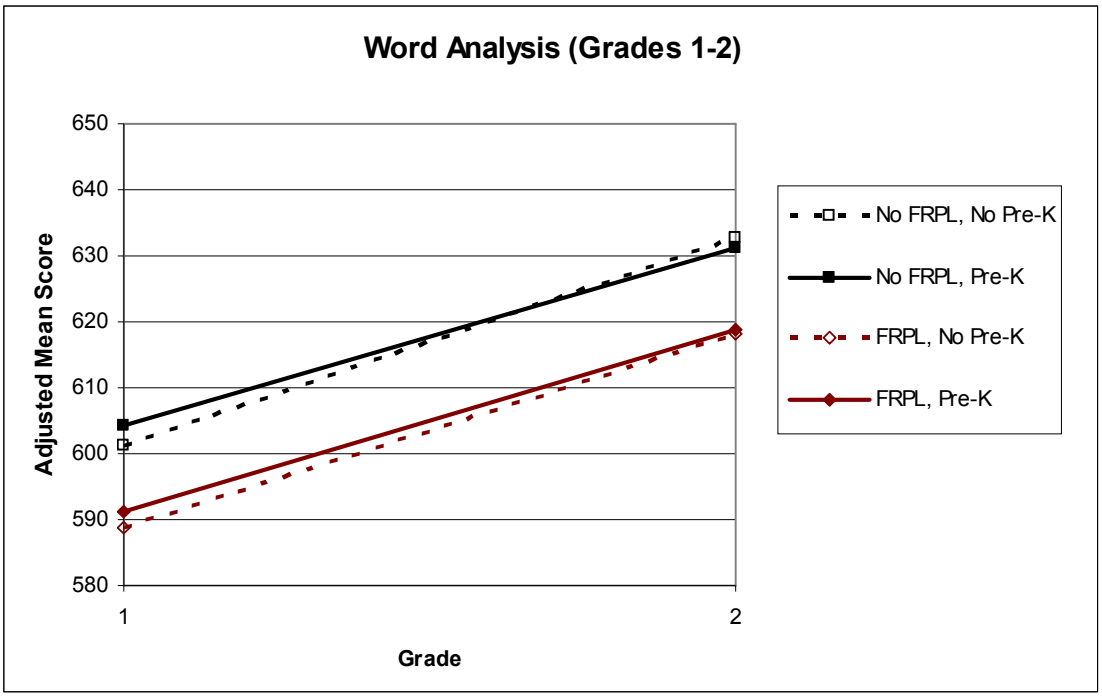


Figure 14. Model-Implied Adjusted Mean Scores for Math Computation in Grades 1-2 for Pre-K and Non-Pre-K Students by Student FRPL Status

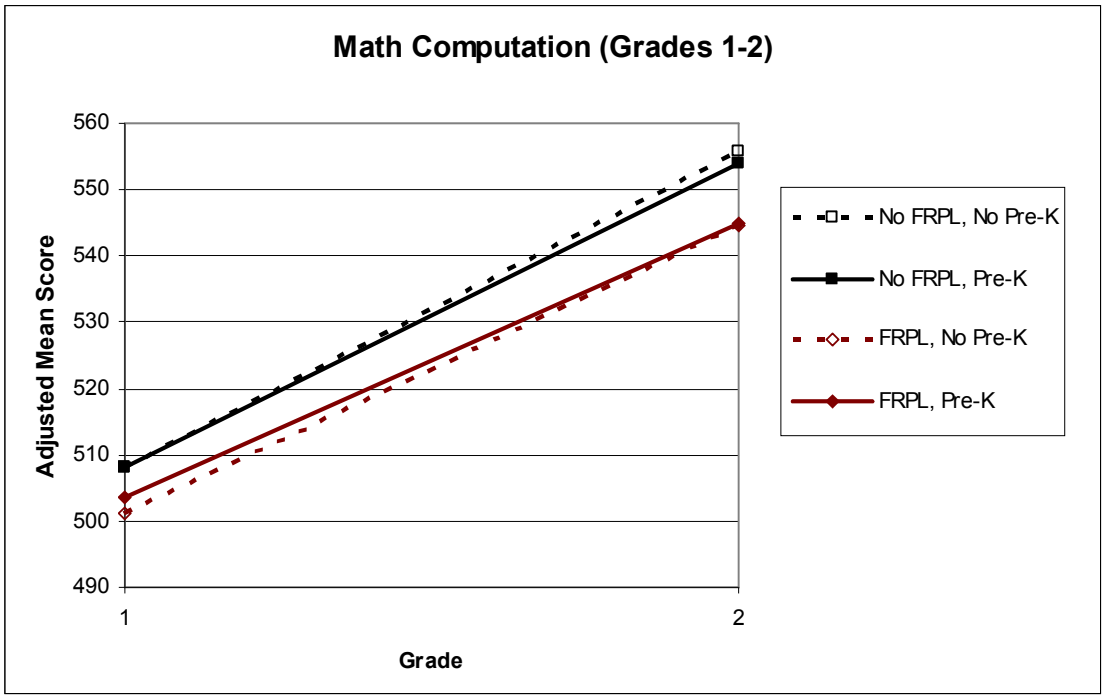


Figure 15. Model-Implied Adjusted Mean Scores for Social Studies in Grades 1-2 for Pre-K and Non-Pre-K Students by Student FRPL Status

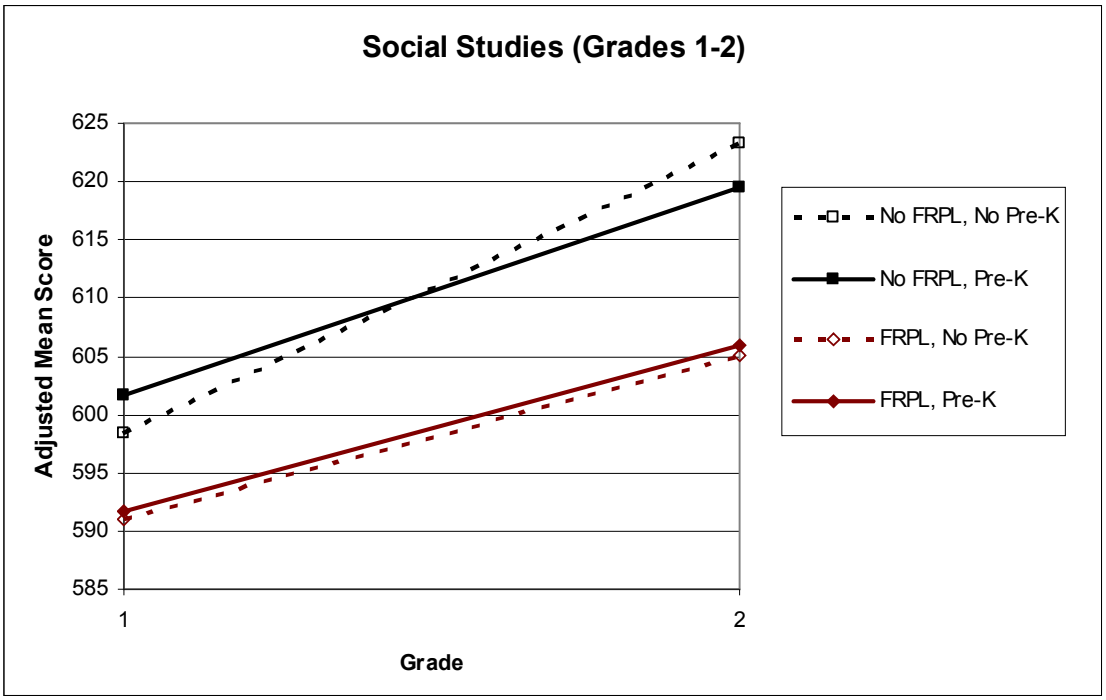
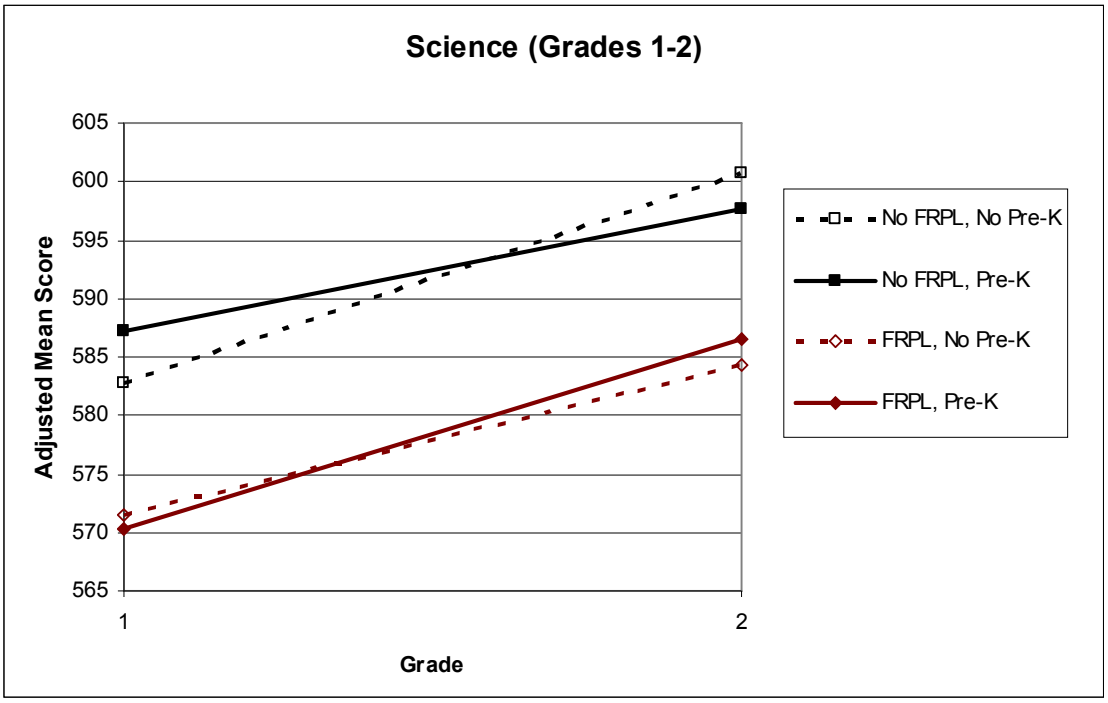


Figure 16. Model-Implied Adjusted Mean Scores for Science in Grades 1-2 for Pre-K and Non-Pre-K Students by Student FRPL Status



The Spelling assessment is only administered in the Second Grade. So, results for Spelling were not examined longitudinally. However, the same basic pattern in results was observed for this assessment as with the other Second Grade assessments—there was no statistically significant difference in student scores attributable to Pre-K participation. Although Pre-K student scores were, on average, slightly higher than their non-Pre-K peers, these differences were not statistically significant after corrections were made for multiple comparisons.

Table 20 summarizes the adjusted mean scores for each group, and Figure 9 illustrates the pattern of results for Second Grade spelling.

Table 20. Model-Implied Adjusted Mean Scores for Pre-K and Non-Pre-K Students—Second Grade

	Group					
Assessment	**Pre-K**	**Non-Pre-K**	**Pre-K FRPL**	**Non-Pre-K FRPL**	**Pre-K No FRPL**	**Non-Pre-K No FRPL**
Second Grade Spelling	588.29	585.29	579.02	575.39	597.55	595.18

Note: Model based on a sample size of 3,734 children.

* Denotes a statistically significant difference between Pre-K and non-Pre-K students, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

** Denotes a statistically significant difference between Pre-K and non-Pre-K students who received FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

*** Denotes a statistically significant difference between Pre-K and non-Pre-K students who did not receive FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

Figure 17. Model-Implied Adjusted Mean Scores for Spelling in Grade 2 for Pre-K and Non-Pre-K Students

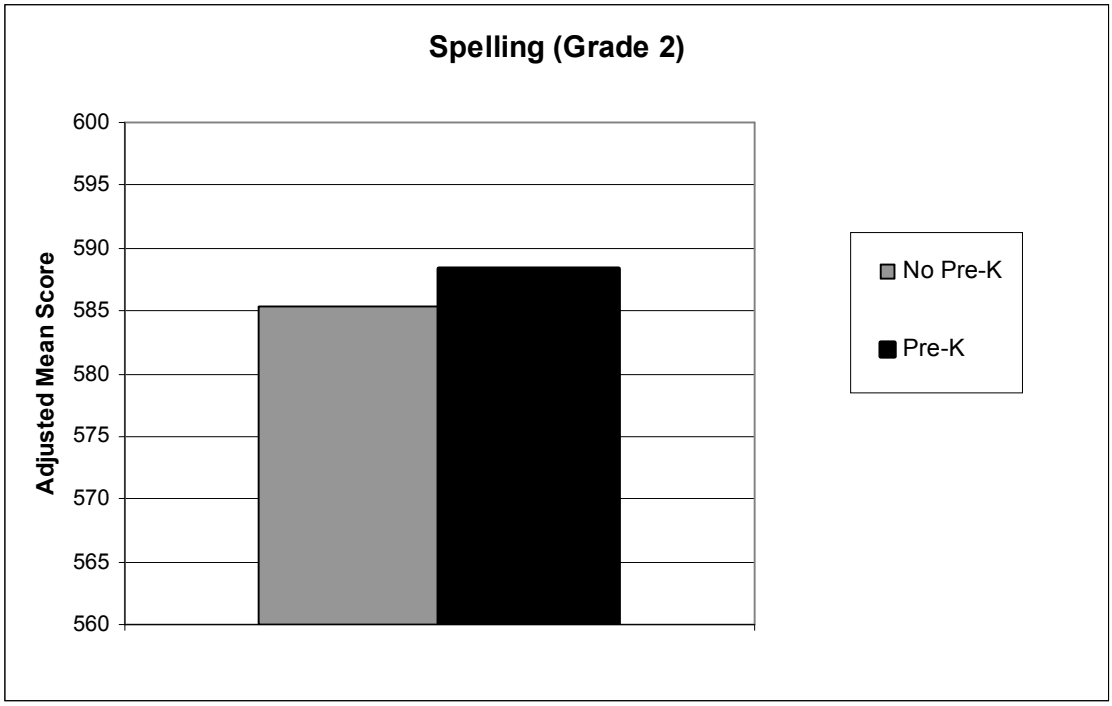
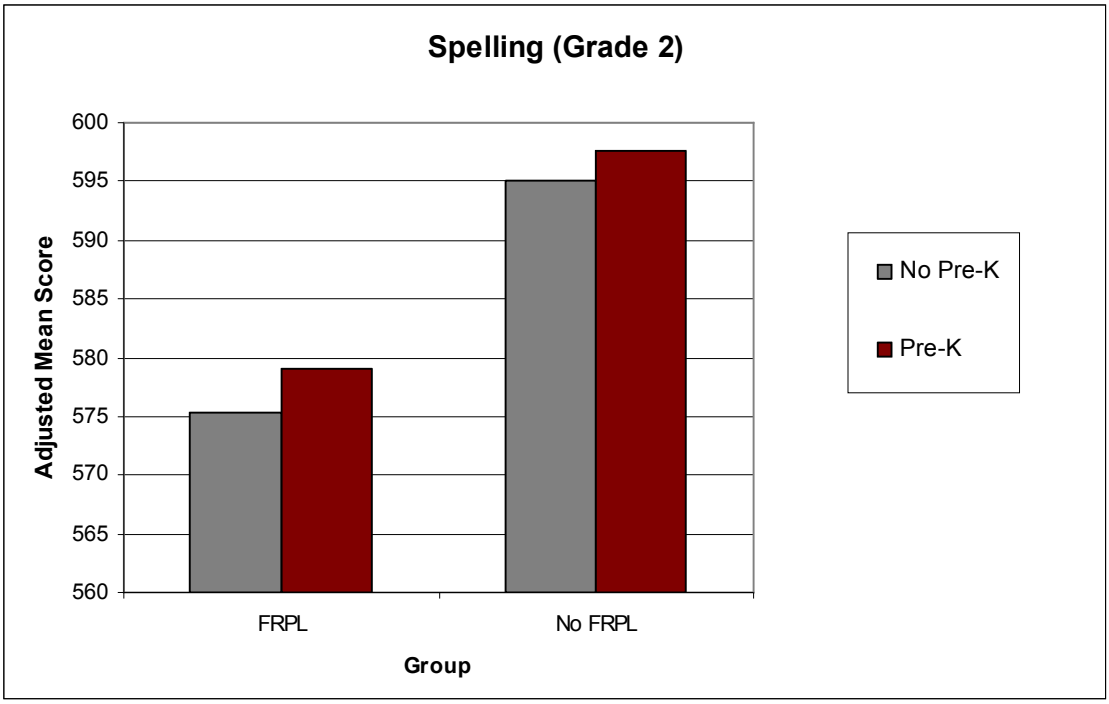


Figure 18. Model-Implied Adjusted Mean Scores for Spelling in Grade 2 for Pre-K and Non-Pre-K Students by Student FRPL Status



Summary: Short-Term Effects

A consistent pattern of results was observed across the Norm-Referenced Assessments administered in Grades K-2. For those assessments administered in Kindergarten, Pre-K students scored better than non-Pre-K students, although there was also a significant difference depending on whether or not a student received FRPL (an indicator of student socioeconomic status). Specifically, among students who received FRPL, students who attended Pre-K scored higher than students who did not attend Pre-K when tested in Kindergarten. Further, among students who did not receive FRPL, students who attended Pre-K scored higher than students who did not attend Pre-K on Kindergarten assessments. However, analysis of assessments administered in First and Second Grade indicate that Pre-K participation was not associated with a statistically significant difference in assessment scores, and in fact the Pre-K and non-Pre-K groups tend to converge by Second Grade.

Criterion-Referenced Assessments (Grades 3-5)

Criterion-Referenced Assessments in Reading, Mathematics, Social Studies, and Science are administered in Grades 3-5. Student performance on these assessments is compared to a predetermined standard (i.e., "cut point") to determine proficiency. The cut points established by TDOE for each of these subjects in each grade are presented in Table 21.

Table 21. TCAP Cut Scores for Reading, Mathematics, Social Studies and Science in Grades 3-5

Final Cut Scores Established in 2004			
Content Area	**Grade**	***Proficient***	***Advanced***
Reading	3	455	496
	4	461	510
	5	467	522
Mathematics	3	448	484
	4	457	507
	5	463	517
Social Studies	3	188	212
	4	190	216
	5	194	217
Science	3	188	213
	4	189	215
	5	191	218

Source: Tennessee Department of Education

Following the pattern observed in other assessments, for Reading, Social Studies, and Science, Pre-K students did not significantly differ from Non-Pre-K students when assessed in the Third Grade. This pattern persisted in Fifth Grade.

However, scores for the Mathematics assessment did reveal one significant effect of Pre-K in the Third Grade, such that among students who received FRPL, students who participated in Pre-K scored, on average, higher than students who did not participate in Pre-K. There was not a statistically significant difference in the Fifth Grade for Mathematics, and there were no other

statistically significant differences indicating that Pre-K participants scored higher relative to their non-Pre-K peers. Table 22 summarizes the adjusted mean scores for the assessments analyzed in Grades 3-5, and Figures 19-22 present the results of the growth curve models for each assessment (Reading, Mathematics, Social Studies, and Science) for the Pre-K and non-Pre-K groups overall, and Figures 23-26 present the results including student FRPL status.

Table 22. Model-Implied Adjusted Mean Scores for Pre-K and Non-Pre-K Students at First and Last Observation (Grades Three and Five)

	Group					
Assessment	**Pre-K**	**Non-Pre-K**	**Pre-K FRPL**	**Non-Pre-K FRPL**	**Pre-K No FRPL**	**Non-Pre-K No FRPL**
Third Grade Reading	485.51	486.68	481.82	480.38	489.20	492.97
Third Grade Mathematics	476.50	475.49	**473.59****	**470.85****	479.41	480.13
Third Grade Social Studies	201.07	201.11	197.83	197.29	204.30	204.93
Third Grade Science	203.22	203.49	200.62	200.01	205.82	206.96
Fifth Grade Reading	505.37	509.03	500.32	501.83	510.42	516.22
Fifth Grade Mathematics	510.96	514.42	509.33	507.21	512.59	521.63
Fifth Grade Social Studies	207.20	208.57	206.54	206.54	207.85	210.59
Fifth Grade Science	204.88	203.95	200.89	200.30	208.86	207.59

Note: Models based on a minimum sample size of 4,320 children.

* Denotes a statistically significant difference between Pre-K and non-Pre-K students, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

** Denotes a statistically significant difference between Pre-K and non-Pre-K students who received FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

*** Denotes a statistically significant difference between Pre-K and non-Pre-K students who did not receive FRPL, at the $p < .05$ level after adjusting for multiple comparisons (means appear in bold).

Figure 19. Model-Implied Adjusted Mean Scores for Reading in Grades 3-5 for Pre-K and Non-Pre-K Students

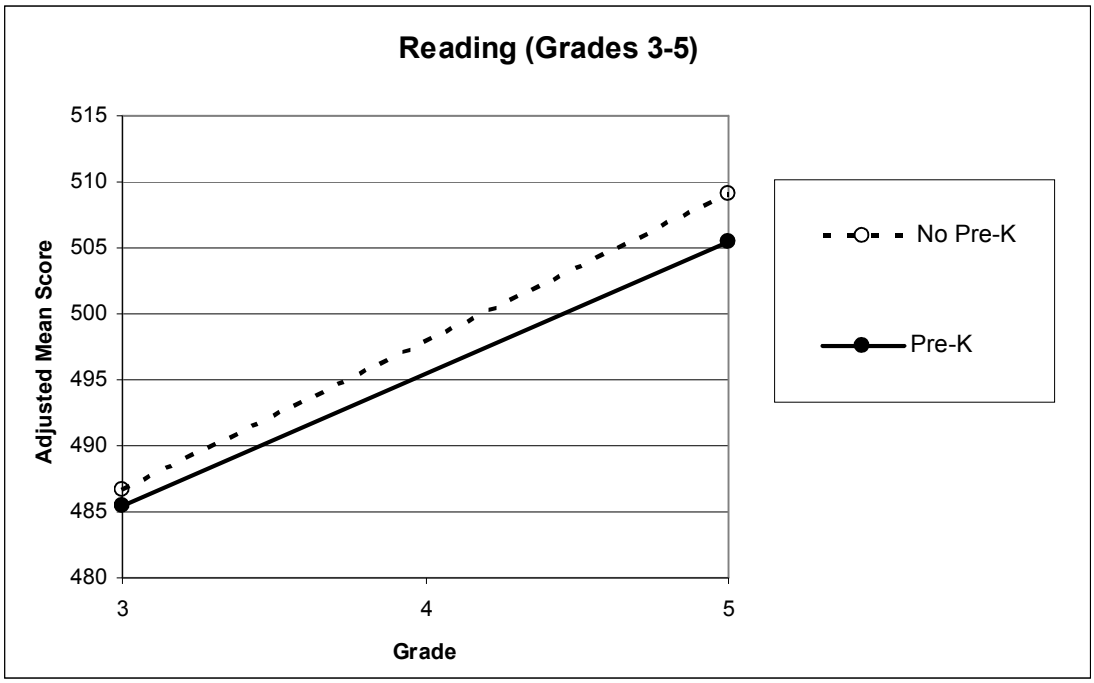


Figure 20. Model-Implied Adjusted Mean Scores for Mathematics in Grades 3-5 for Pre-K and Non-Pre-K Students

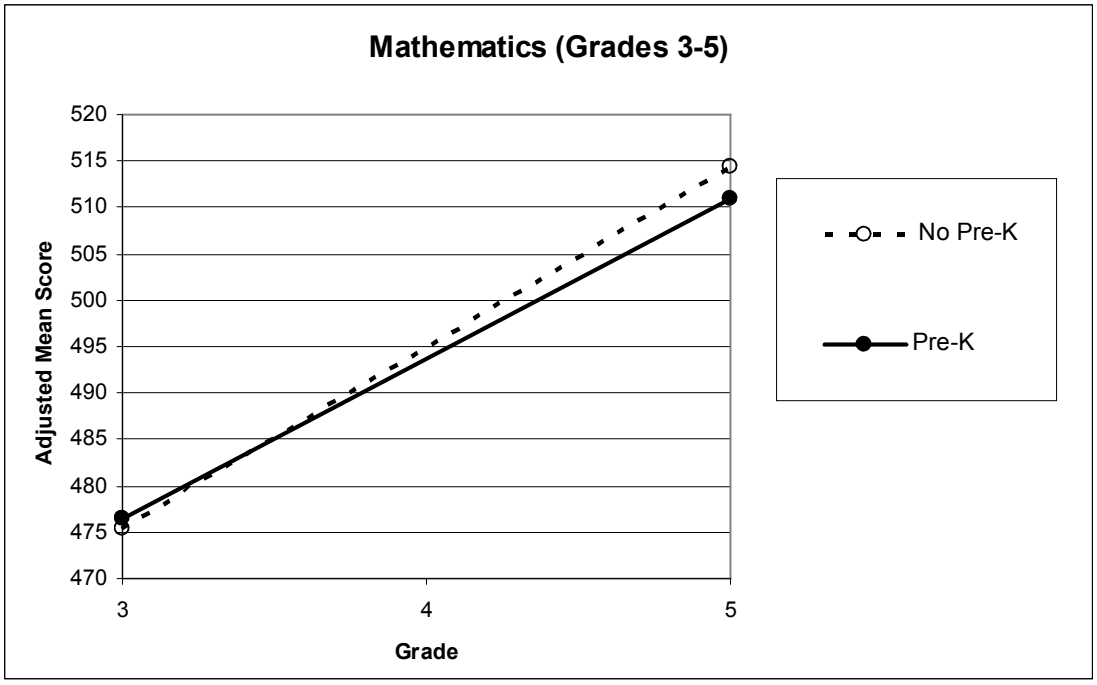


Figure 21. Model-Implied Adjusted Mean Scores for Science in Grades 3-5 for Pre-K and Non-Pre-K Students

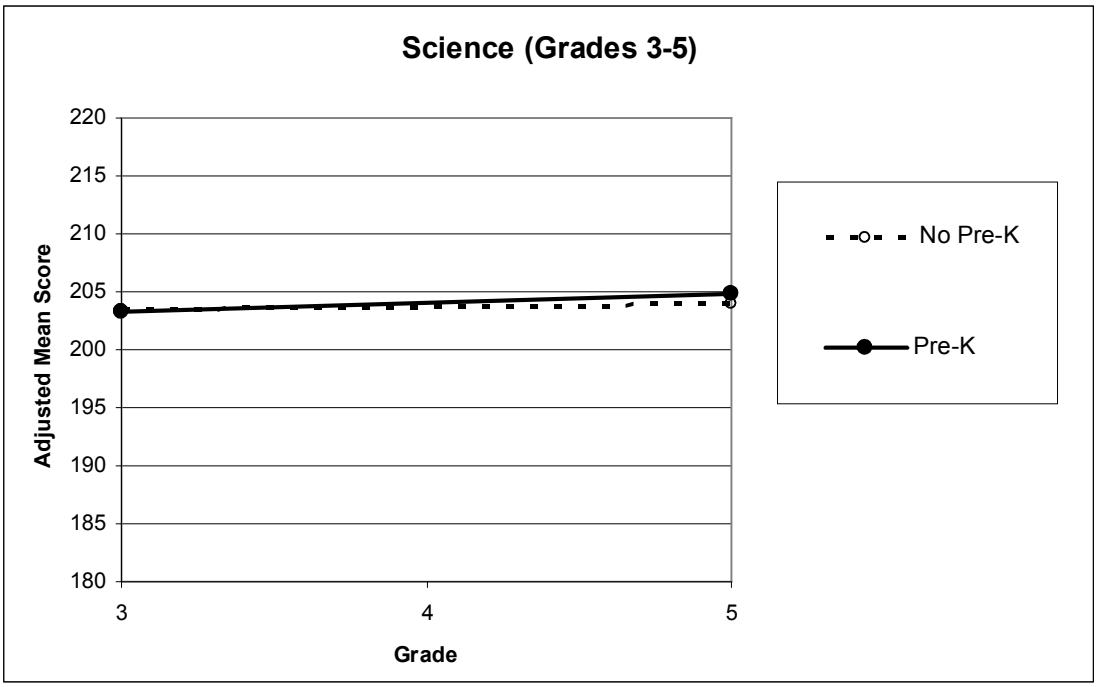


Figure 22. Model-Implied Adjusted Mean Scores for Social Studies in Grades 3-5 for Pre-K and Non-Pre-K Students

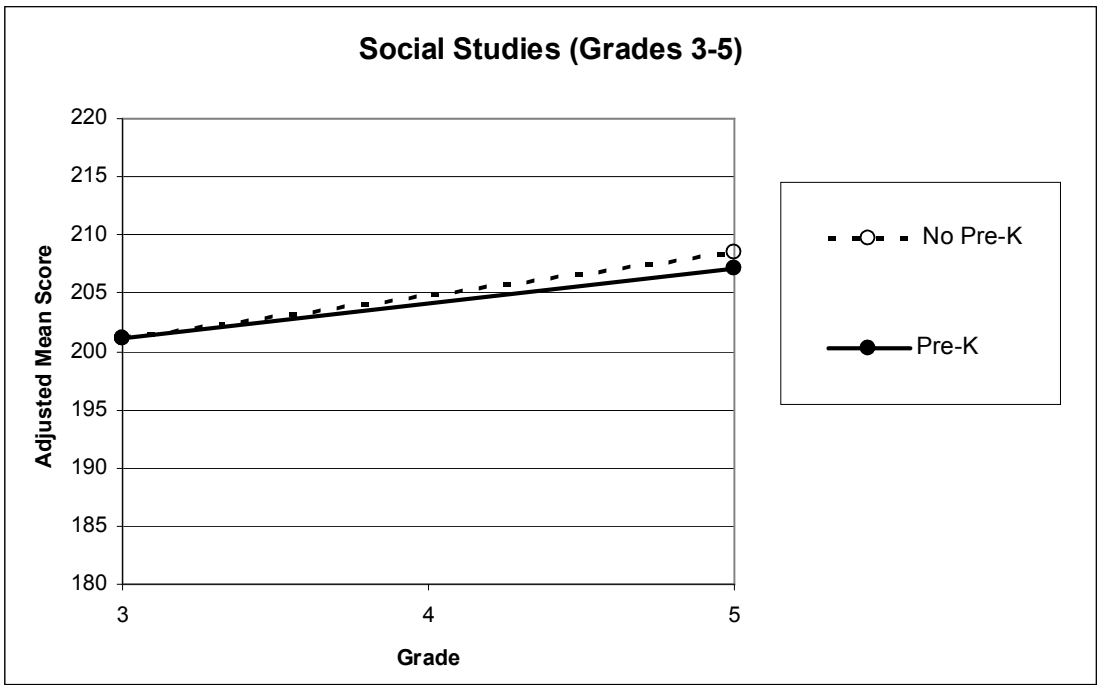


Figure 23. Model-Implied Adjusted Mean Scores for Reading in Grades 3-5 for Pre-K and Non-Pre-K Students by Student FRPL Status

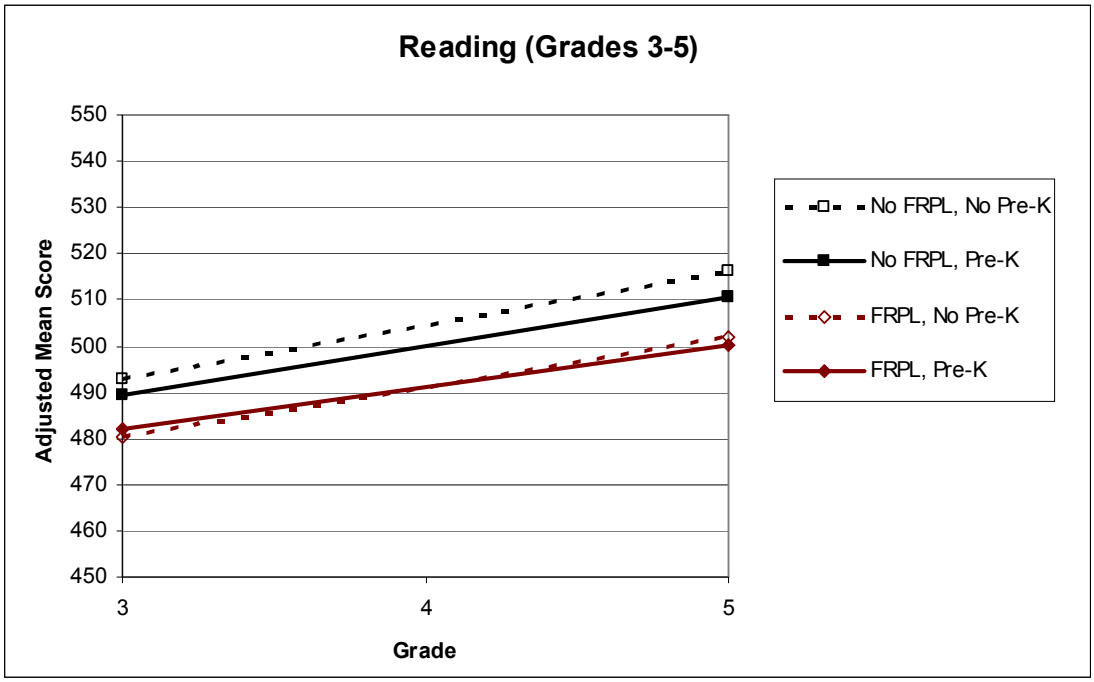


Figure 24. Model-Implied Adjusted Mean Scores for Mathematics in Grades 3-5 for Pre-K and Non-Pre-K Students by Student FRPL Status

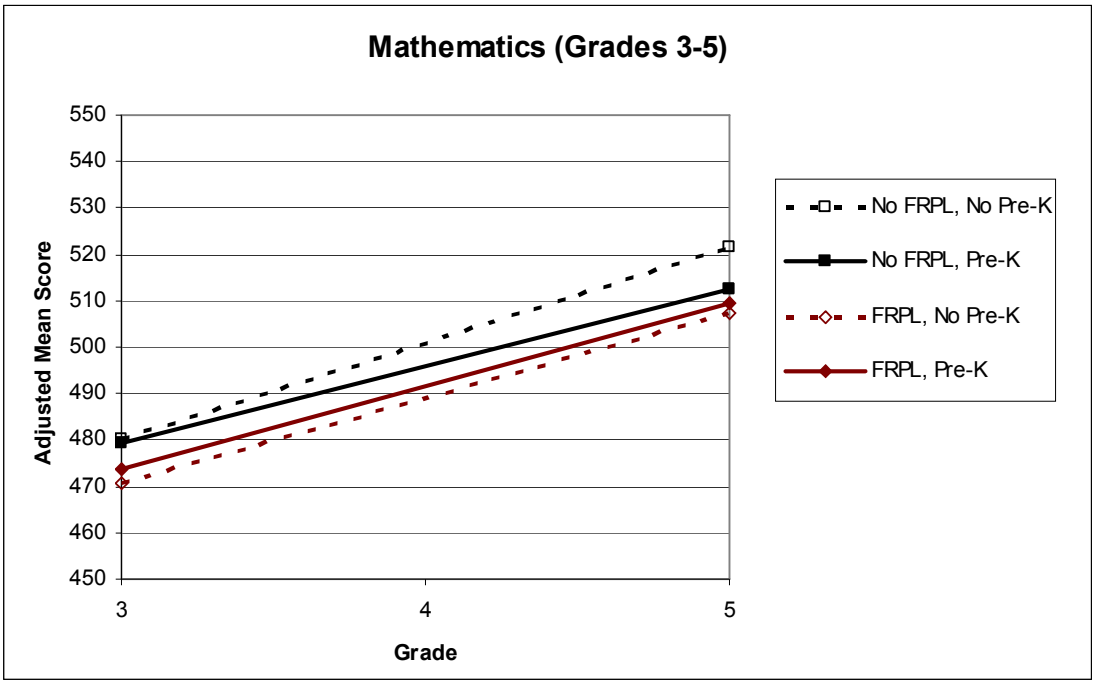


Figure 25. Model-Implied Adjusted Mean Scores for Science in Grades 3-5 for Pre-K and Non-Pre-K Students by Student FRPL Status

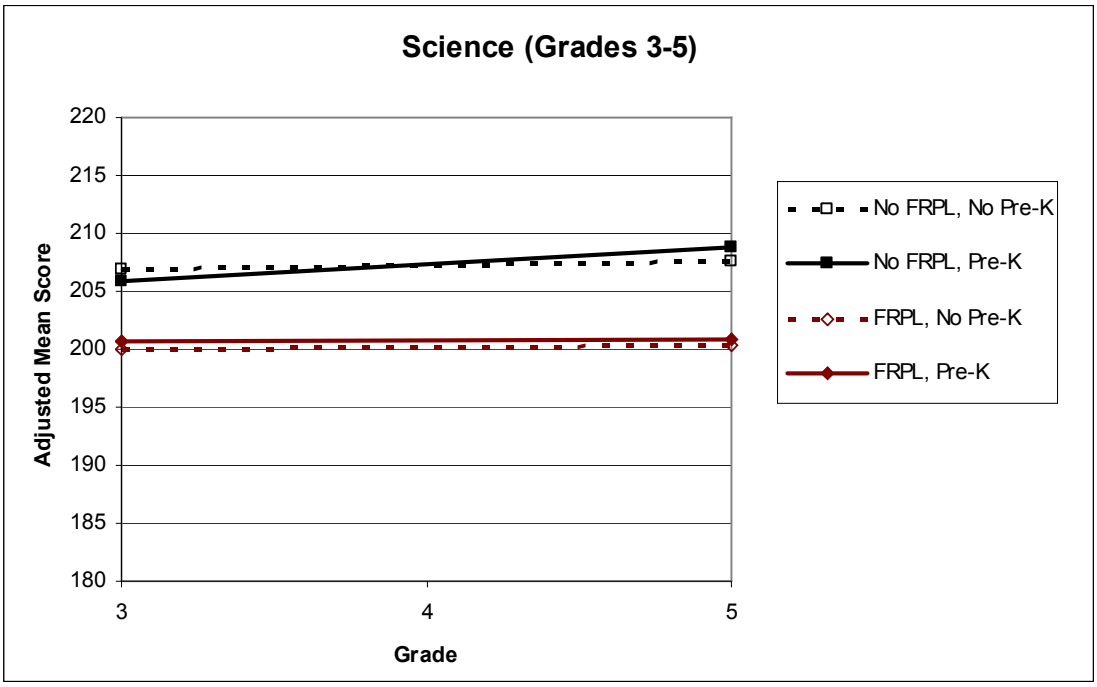
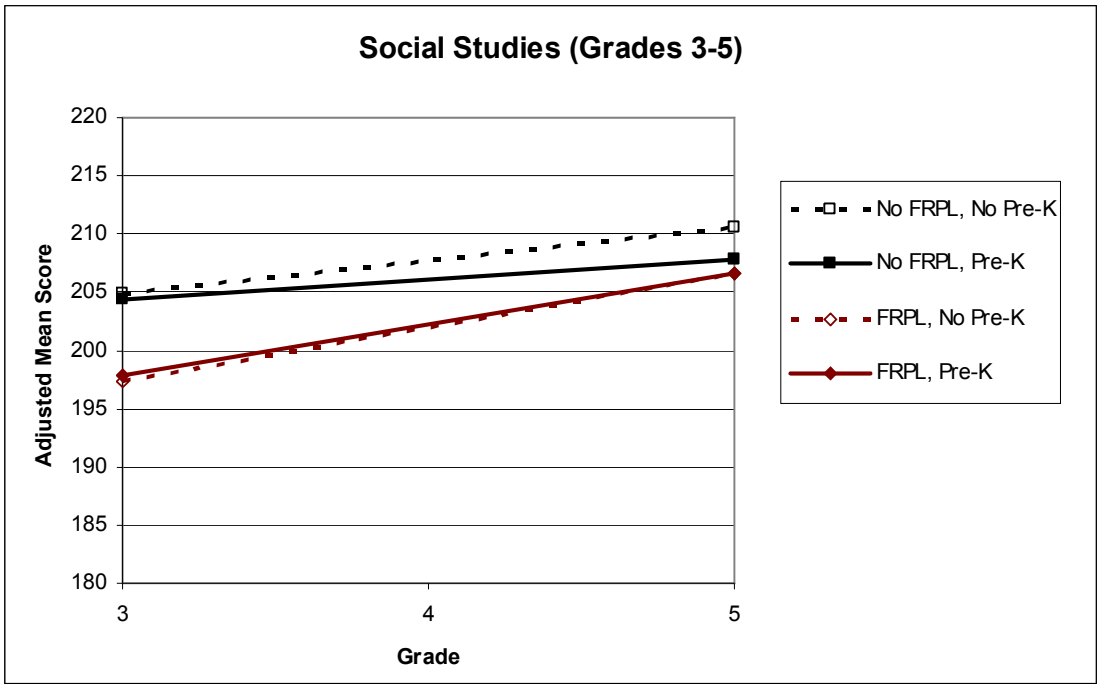


Figure 26. Model-Implied Adjusted Mean Scores for Social Studies in Grades 3-5 for Pre-K and Non-Pre-K Students by Student FRPL Status



Summary: Long-Term Effects

A consistent pattern was observed across the results of Criterion-Referenced Assessments administered in Grades 3-5, and this indicated that Pre-K students performed similarly to non-Pre-K participants on assessments conducted in these grades. Growth curve models showed only one statistically significant difference associated with Pre-K participation on one assessment in one grade level. Specifically, in the Third Grade, students who had participated in Pre-K performed slightly better on the Mathematics assessment relative to non-Pre-K students, yet this was a modest difference of only 3 points, on average. There were no other statistically significant differences between the Pre-K and non-Pre-K groups among students who received FRPL or those who did not.

Additional Effects

Analysis of fixed effects in the models examined the unique effects of gender, race, absences, special education, retention, and native language and their interactions with Pre-K participation and FRPL status. These were explored for all models (one, two, and three time points) and for all assessments. There were no consistent interactions between Pre-K participation and gender, race, or any of the other predictor variables. This can be interpreted to indicate that overall, the general pattern of short- and long-term results discussed above holds for all subgroups of students—in other words, the general pattern observed is the same for both male and female students, white and non-white students, and so on.

Although exploratory analyses conducted in the First Interim Report suggested that Pre-K participation interacted with race and gender, it is important to note that FRPL status was not taken into consideration as a predictor variable in those analyses. Including FRPL status and exploring its effects in the context of these other predictors reveals that FRPL status is a significant predictor of these student outcomes, and once this variable is accounted for in the model, the impact of these other variables (e.g., race, gender, etc.) is rendered nonsignificant.

General Summary

Pre-K participation appears to be associated with significant differences in Kindergarten assessments of Reading, Language Arts, and Mathematics such that Pre-K participants score higher than a matched sample of non-Pre-K participants, although students' socioeconomic status (i.e., whether they receive FRPL) also plays a significant role in their outcomes on these assessments. The models employed in the present study did not find that this relative advantage persisted over time, however, and as students moved through higher grades, their scores tended to converge. Specifically, Pre-K and non-Pre-K students performed similarly overall. Further, students receiving FRPL tended to be more similar to one another irrespective of their participation in Pre-K, and students who did not receive FRPL tended to be more similar to one another again, irrespective of Pre-K participation. On the whole, the results demonstrate a clear pattern of convergence, despite initial differences associated with Pre-K participation on assessments conducted in Kindergarten.

Further analysis is warranted to explore these results further, as discussed in the next section. In addition, we wish to address the fact that the statistical control variables used to examine short- and long-term effects associated with Pre-K participation in the present study (retention, attendance, and special education measured after the Pre-K year) are variables which themselves could have been affected by participation in the Pre-K program. These variables were included in the present model due to their theoretical significance. In deciding whether or not to include these variables in the models we did take into consideration the fact that including these controls may have the additional effect of "controlling out" some of the Pre-K program

effect. However, failing to account for these important potential sources of variability in students' scores could have led to inaccurate (i.e., biased) results, thereby hampering our ability to provide an accurate reflection of student progress—regardless of Pre-K experience. The possible relationship between Pre-K experience and predictors of academic success such as retention and attendance has not gone unnoticed. However, further exploration is required before a definitive picture of this relationship can be presented. This will be examined in future reports.

Next Steps and Subsequent Reports

The Second Interim Report represents an additional step in the process of conducting a comprehensive analysis of student outcomes between the years 1998-1999 and 2008-2009. Table 23 summarizes the cohorts of Pre-K participants and the school years covered in subsequent reports.

Table 23. Pre-K Cohorts and School Years Covered in the Schedule of Reports

	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009
Cohort 1	Pre-K	K	1st	2nd	3rd	4th	5th				
Cohort 2		Pre-K	K	1st	2nd	3rd	4th	5th			
Cohort 3			Pre-K	K	1st	2nd	3rd	4th	5th		
Cohort 4				Pre-K	K	1st	2nd	3rd	4th	5th	
Cohort 5					Pre-K	K	1st	2nd	3rd	4th	5th
Cohort 6						Pre-K	K	1st	2nd	3rd	4th
Cohort 7							Pre-K	K	1st	2nd	3rd
Cohort 8								Pre-K	K	1st	2nd
Cohort 9									Pre-K	K	1st
Cohort 10										Pre-K	K
Cohort 11											Pre-K
		First Interim Report					Second Interim Report			2008 Annual Report Third Interim Report	2009 Annual Report Final Report

The next report will be an Annual Report (due in Fall 2008).

A key advantage of this next report, and the reports that follow, is that Pre-K participation will be identified by the EIS rather than the Pre-K demographic file. As mentioned previously, starting with the 2006-2007 school year, the Tennessee EIS began including data for Pre-K students. Thus, the next report will include the first cohort of Kindergarten students who attended Pre-K in 2006-2007. Because the EIS provides more complete coverage of information than the Pre-K demographic file, we will be able to further reduce instances of missing demographic data for

students. Along similar lines, there are fewer cases of missing unique student identifiers (i.e., Social Security Number, provided to us in encrypted form) in the EIS as compared to the Pre-K demographic file. As mentioned previously, students for whom no Social Security Number was provided cannot be included in the analyses.

Further, although the Pre-K demographic file will continue to be used to identify Pre-K students from previous years, the availability of additional years of assessment and EIS data will allow us to cross-check the records of more Pre-K students who could not be included in previous analyses because their records contained questionable or inconsistent information.

With each successive report, we gain additional and more complete information about each cohort of students, and move closer to the point where we can include all cohorts over all years to have the most accurate vantage point on change over time for students who participated in Pre-K. In the present report, we were able to capitalize on additional student-level information not available for the First Interim Report, and thus were able to explore student outcomes in a more comprehensive manner. It will be important to see whether the same patterns observed here persist over time.

In addition, given the fact that the only reliable significant difference between the Pre-K and non-Pre-K group was found in Kindergarten assessments, there are several important issues that we feel merit additional investigation. First, given that only a small number of schools administer assessments in Kindergarten, a logical next step is to investigate further the characteristics of the schools and how these school-level characteristics relate to Pre-K students' outcomes. This would include characteristics such as school location and general character (e.g. school achievement levels) and to understand better what parts of the Tennessee school system are represented. Thus, additional analyses are planned to take these factors into consideration.

An outstanding research question involves the relative impact of various program characteristics on student outcomes—for example, are specific characteristics of individual Pre-K programs more or less likely to be associated with beneficial outcomes? The data sources provided for analysis provide very little descriptive information on individual programs, and based on information provided about these programs to date, many program characteristics are fixed, meaning that there is very little variability in terms of teacher qualifications, classroom size, and so on. This makes such an analysis difficult, if not impossible, to conduct. Thus, more discussion with the Office of the Comptroller and the Office of Early Learning is necessary to identify appropriate variables to include in a program-level analysis as well as to identify data sources for these variables.

Finally, we understand that although this report is able to answer with some confidence the primary research questions of interest for the present evaluation, there are many outstanding questions about the impact and effectiveness of the Pre-K program that remain unanswered—and indeed, this report may generate some new questions. We wish to conclude the present interim report with the caveat that no single study can address every possible question about a program as large as Tennessee's Pre-K program (and one with many passionate proponents), particularly when the data are limited and the methodology is retrospective. However, we have used the best possible methods for the data available to provide as accurate a picture as possible of the performance of Pre-K students, and will continue to explore this data as well as additional years' data as the evaluation progresses.

Appendix A. Background and Implementation of Tennessee's Pre-Kindergarten Program

Across the nation, access to high-quality state-funded Pre-Kindergarten (Pre-K) has steadily increased in the last 10 years. The National Institute for Early Education Research (NIEER) estimates that almost 1 million children participated in state Pre-K initiatives in 2005-2006, and spending in the states offering Pre-K totaled over $3 billion.[21] Although state standards vary widely, more than three-quarters of state programs adhere to comprehensive early learning standards and more than half require teachers to have a Bachelor's degree; 73% require teachers to have specialized Pre-K training. NIEER estimates 14% of 4-year-olds participated in state-funded Pre-K in 2002, but in 2006 20% of 4-year olds were enrolled.

The State of Tennessee has been funding Early Childhood Education (ECE) since the 1990s. Legislation enacted in 1996 permitted the creation of Pilot early childhood and Pre-Kindergarten programs for economically disadvantaged three- and four-year-olds. In the 1998-1999 school year, 30 Pilot Pre-K classrooms were created, serving approximately 600 students. Since then the program has grown to over 934 classrooms, serving approximately 17,000 children. Table A1 summarizes the number of students served and the number of classrooms in operation in Tennessee since 1998-1999, according to Tennessee Department of Education, Office of Early Learning.

Table A1. Number of Students Enrolled in Tennessee Pre-K, 1998-1999 to 2007-2008

Program Year	Students Served	Number of Classrooms
1998-1999	600	30
1999-2000	600	30
2000-2001	3,000	150
2001-2002	3,000	90
2002-2003	3,000	150
2003-2004	2,900	150
2004-2005	2,900	147
2005-2006	8,900	446
2006-2007	13,000	677
2007-2008	17,308	934

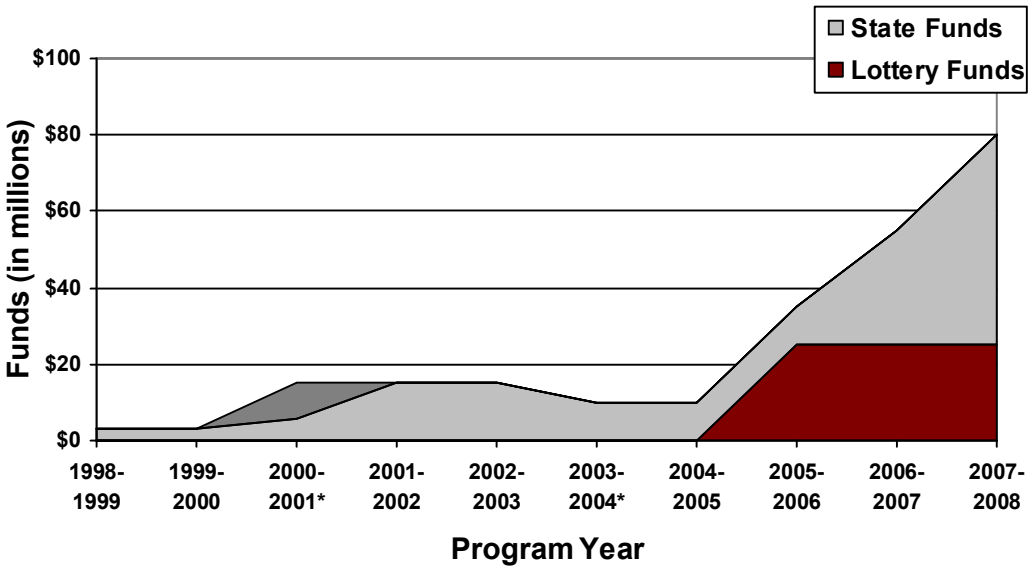
Source: State of Tennessee, Office of Early Learning

The state Pre-K program has benefited from strong support from the Governor and bipartisan support in the Tennessee General Assembly. Together they passed the Voluntary Pre-K for Tennessee Act of 2005, increasing the state's investment in Early Childhood Education and access for four-year-olds. The state allocated $25 million from the excess net education lottery proceeds to fund approximately 300 new Pre-Kindergarten classrooms for at-risk four-year-olds—effectively tripling the number of students served. In the past three program years, state

[21] Barnett, W.S., Hustedt, J.T., Hawkinson, L.E., Robin, K.B. (2006). *The State of Preschool 2006*. National Institute for Early Education Research. Downloaded from the Internet at http://nieer.org/yearbook/.

contributions have surged to $55 million for a total of $80 million for 2007-2008. Figure A1 shows the trends in funding over 10 program years.

Figure A1. Tennessee Pre-K Funding by Source, 1999-2008



Notes: $9 million in funding was awarded and disseminated mid-year in program year 2000-2001 (dark gray area). TANF funding ended in 2003-2004 school year.

Pre-K Pilot Sites

The expansion of the Pre-Kindergarten program in 2005 resulted in two systems of Pre-Kindergarten instruction: the Pilot programs that were begun in 1998 and the lottery/general fund-funded programs that were begun in 2005. The two systems are alike in their classroom requirements regarding teacher credentials, class size, and curricular focus; however, they differ in their funding amount and source and their affiliation with the public school system. Pilot Pre-K program sites were not required to be affiliated with a local education agency (LEA) although most were located in schools. Although the majority of Pilot Pre-K providers are LEAs, 14 private providers (approximately 22 classrooms) have continued to serve three- and four-year olds. In 2006-2007, approximately 3,000 students were served in the Pilot program.

In fiscal years 2002 and 2003, funding of the Pilot sites was supplemented by federal Temporary Assistance for Needy Families (TANF) funds. When the TANF funding ended in the 2003-2004 school year, the state became the program's sole funding source and each classroom's funding was reduced by $30,000. In 2005-2006, each classroom received $65,000 in state funds. A local match is not required; however, most pilot sites supplement the state allocation with local funds.

Student Eligibility

Enrollment in the Voluntary Pre-K for Tennessee program is based upon legislation (TCA 49-6-101-104). The Pre-K state statute specifies that each LEA is authorized to and may enroll any at-risk child who is four years old by September 30 and resides in the geographic area served by the LEA. Priority is given to those children who are eligible for the free/reduced price lunch (FRPL) program. The state of Tennessee Department of Education encourages school systems to accurately identify the number of unserved at-risk children in the school district, making every effort to fill Pre-K classrooms with at-risk children.

If, however, there is space available after priority is given to at-risk four-year-olds, the LEA may enroll students with disabilities, students identified as English language learners (ELLs), students in state custody, or students identified as educationally at-risk for failure due to circumstances of abuse or neglect.

At the end of the first Pre-K student attendance period, if an insufficient number of children meeting the aforementioned enrollment requirements are enrolled in a specific classroom, an LEA may submit a written request to the Office of Early Learning for approval to enroll children identified with other at-risk factors as determined by the local school board and the Pre-K Advisory Council, such as children with a parent(s) in the military deployed to active duty, teen parents, or parents with limited education. In these cases a written request must be submitted and approved by the Office of Early Learning.

If not enough at-risk children enroll to fill a classroom, students who do not meet any at-risk criteria but who are considered unserved or underserved may be enrolled after a written request is submitted and approved by the Office of Early Learning.

Classroom Requirements

Classroom requirements for all Pre-K classrooms in Tennessee are the same for all sites, regardless of whether they are considered Pilot or Lottery/general fund-funded Pre-K programs. The requirements are as follows: [22]

- Maximum class size is 20 students.
- Each class must have at least one licensed teacher who is certified in early childhood education, and at least one educational assistant who has either a child development associate credential (CDA) or an associate degree in early childhood education, or who is working toward acquiring these credentials.
- The program must provide a minimum of five and one-half hours of quality instructional time per day.
- Classroom instruction must include the use of an educational, age-appropriate curriculum that aligns with the state department of education approved early learning standards and includes literacy, writing, math, and science skills.
- Instruction must also include a developmental learning program that addresses the cognitive, physical, emotional, social, and communication

[22] Note: Complete scope of services for 2007-2008 (including program requirements) can be found on the State of Tennessee, Department of Education, Office of Early Learning web page: http://state.tn.us/education/prek/documents/VoluntaryScopeofServices2007-08.doc

areas of child development.

- In addition, each program must comply with the state board of education's early childhood education and Pre-Kindergarten program rules and policies.

Program Effectiveness

Clearly, Tennessee's Pre-K program has experienced significant growth over the past three years. In 2005, the program served approximately 3,000 three- and four-year-olds in 148 classrooms funded with $10 million in state revenue, but is expected to increase to over 900 classrooms (17,000 students) in 2007-2008, representing an increase of over 750 classrooms in just 3 years. By 2008, state Pre-K is expected to have expanded to 132 of the 136 school systems in 93 of the 95 counties in Tennessee. The Governor and Legislature have made Pre-K a priority in Tennessee, and funding for Pre-K education has increased from $10 million in 2004-2005 to $80 million in 2007-2008 through excess lottery funds and state revenue.

Collaboration is a distinctive characteristic of Tennessee's Pre-K program, as evidenced by the importance of classroom partnerships in the TN Pre-K program. In 2006-2007 there were 148 collaborative classroom partnerships between 39 local school systems and non-profit and for profit providers. Tennessee statute allows for state collaboration with agencies such as Head Start, Even Start, for-profit and not-for-profit child care providers, faith-based agencies, community-based agencies, and higher education institutions. Further, the program requires the active participation and collaboration of stakeholders at the local and state level in the form of Community and State Pre-K Advisory Committees.

The Tennessee Voluntary Pre-K program has been recognized as achieving 9 out of 10 quality standard benchmarks of the National Institute for Early Education Research (NIEER), for the past two years—one of only 6 states to achieve a score of 9 or 10.[23] The NIEER benchmarks include:

- Comprehensive early learning standards.
- Teacher degree of BA or higher.
- Teacher specialized training in Pre-K.
- Assistant teacher degree of CDA or equivalent.[24]
- Teacher in-service of at least 15 hours per year.
- Maximum class size of 20 students.
- Staff: child ratio 1:10 or better.
- Vision, hearing, health screenings and one support service.
- At least one meal per day.
- Monitoring/site visits.

As the program continues to grow and more children are exposed to high-quality early childhood education in Tennessee, research is increasingly able to investigate the short-

---

[23] Source: *The State of Preschool 2006*. National Institute for Early Education Research, www.nieer.org.

[24] According to NIEER, Tennessee meets all the NIEER criteria with one exception (assistant teacher degrees). The full report is available for download from the Internet at http://nieer.org/yearbook/.

and long-term impact of Pre-K on student outcomes in elementary and middle school. The state of Tennessee has been collecting data on student participation in Pre-K since the inception of the Pilot Pre-K program in 1998, and is in the unique position to track student outcomes longitudinally.

A comprehensive program overview is available from the Tennessee Alliance for Early Education. This document, "Voluntary Pre-K in Tennessee, Understanding the Collaboration Model" (March, 2008), is available for download through the State of Tennessee Office of Early Education:
http://www.tennessee.gov/education/prek/doc/Vol_prek_under_collab.pdf

Appendix B. Technical Specification of Models

The models presented in this report can be understood through a general 3-level hierarchical linear model that accounts for either 1) time (i.e., observation) nested within child and child nested within school or 2) child nested within school and school within school district. The general model is presented relying heavily on the Raudenbush and Bryk (2002) terminology. The general model is presented in "levels" and is discussed in terms of multiple observations within child (i.e., a growth model) and multiple children within school. Variations in parameter use and interpretation are discussed in the Specific Model section.

Level 1

Level 1 defines the relationship between grade and the outcome of interest:

$$y_{tis} = \pi_{0is} + \pi_{1is}(grade_t) + e_{tis}, \tag{1}$$

where

$$e_{tis} \sim N(0, \sigma^2). \tag{2}$$

In Equation 1, y_{tis} denotes outcome y at time t for individual i in school s. The score is defined by an intercept, π_{0is} (individual i's level of y when $grade_t$ equals 0), and a slope, π_{1is} (individual i's model implied mean increase in y given a 1 year increase in $grade_t$). The residual, e_{tis}, captures the time-specific deviation from the observed score for individual i in school s. This deviation is the "error" in prediction not otherwise accounted for by unique individual or school variability. As described in Equation 2, e_{tis} is assumed to be normally distributed with a mean of 0 and a variance of σ^2.

e_{tis} is not the only variance component in the general model. Indeed, the intercept and slope are "random" coefficients allowed to vary over individual thus allowing children to have unique scores within and over time. This unique individual variability is parameterized in Level 2 of the general model.

Level 2

Level 1 parameters π_{0is} and π_{1is} (individual starting point and rate of growth over time) are the outcomes of interest in Level 2 of the general model:

$$\pi_{0is} = \beta_{00s} + \sum_{q=1}^{Q_0} (\beta_{0qs} X_{qi}) + r_{0is}$$
$$\pi_{1is} = \beta_{10s} + \sum_{q=1}^{Q_1} (\beta_{1qs} X_{qi}) + r_{1is}, \quad (3)$$

where

$$r_{pis} \sim \mathrm{N}\left(0, \begin{bmatrix} \tau_{00} & \\ \tau_{10} & \tau_{11} \end{bmatrix}\right) . \quad (4)$$

Equation 3 states that an individual's initial score in school *s* (i.e., π_{0is}) is a linear combination of the overall mean score within school *s*, β_{00s}, the sum of the relationship between the initial score and *q* predictors (q = 1, 2, … Q_p for p coefficients) and an individual deviation, r_{0is}, from the mean score within school *s* after controlling for *q* predictors. An individual's unique rate of change is modeled in a similar fashion; see Equation 3. More specifically, an individual's rate of change over grade (i.e., π_{1is}) is a linear combination of the overall mean rate of change within school *s*, β_{10s}, the sum of the relationship between the rate of change and *q* predictors and an individual deviation, r_{1is}, from the mean rate of change within school *s* after controlling for *q* predictors. Individual deviations from the mean initial score and rate of change are assumed to be bivariate normally distributed with covariance τ_{10}. The variances of the intercept (τ_{00}) and slope (τ_{11}) were further decomposed at the school level.

Level 3

Level 3 defines the Level 2 parameters (β_{00s} and β_{pqs}) as outcomes of interest such that

$$\beta_{00s} = \gamma_{000} + u_{00s}$$
$$\beta_{10s} = \gamma_{100}$$
$$\beta_{pqs} = \gamma_{pq0}, \quad (5)$$

where *p* = 0 when predicting β_{00s}, *p* = 1 then predicting β_{10s}, q = 1 to Q_p, and

$$u_{00s} \sim N(0, \tau_{000}) . \quad (6)$$

Each of the 2 + *q* x *p* Level 2 parameters is defined by an intercept. These intercepts denote the relationship between the outcome and a predictor averaged across all schools. γ_{000} is unique because, when $grade_t$ equals 0, γ_{000} denotes the initial mean outcome score over all individuals

in all schools. Also note, the inclusion of u_{00s} in Equation 5 allows individual schools to deviate from the overall mean score γ_{000}. As can be seen in Equation 6, this residual is assumed to be normally distributed with a mean of 0 and a variance of τ_{000}.

General Model

Given the parameterizations for each level outlined above the general model in its reduced form (i.e., substituting and combining terms) is:

$$y_{tis} = \gamma_{000} + \sum_{q=1}^{Q_0} (\gamma_{0qs} * X_{qi}) + \sum_{q=0}^{Q_1} (\gamma_{1qs} * X_{qi} * grade_t) + [e_{tis} + r_{0is} + r_{1is}(grade_t) + u_{00s}]. \quad (7)$$

All deviations are distributed as described in Equations 2, 4, and 6. As will be demonstrated below, specific models presented in this report are nested within the general model found in Equation 7.

Specific Models

Three models were outlined in the text of this report: the single time point model, the difference model, and the growth model. This section outlines how each of these models is related to the general model presented in Equation 7. We begin with the most complex model, the growth model, moving then to the difference model, and finally the least complex model, the single time point model.

<u>Growth Models</u>

While the growth models presented in this report are the most complex of the models used, the general model was developed to be consistent with a growth model. Therefore, Equation 7 and the text leading up to Equation 7 describe the growth model in its entirety. It is important to note that while estimating a growth model, as it is presented in Equation 7, requires the estimation of five random effects (σ^2, τ_{00}, τ_{10}, τ_{11}, and τ_{000}), for a given outcome. If supported by the data, one or more of these effects may be constrained to zero.

Difference Models

The difference models are parameterized just as the growth models with one exception: two of the random effects (τ_{10} and τ_{11}) are constrained to zero in all difference models. Given that all outcomes examined with difference models are measured on only two occasions, there was little information in the data to estimate individual variability in change over time and thus no reason to estimate the covariance between the intercept and change parameter. All other model parameters are interpreted as described above. As with the growth models, for a given outcome, if supported by the data, one or more additional random effects may be constrained to zero.

Single Time Point Models

The single time point models are the simplest models in this report to parameterize but require the greatest degree of modification to the general model in Equation 7. As with the difference models, τ_{10} and τ_{11} are constrained to zero. However, because the outcomes of interest were only measured within a single grade, $grade_t$ can also be constrained to 0. Constraining these parameters results in:

$$y_{tis} = \gamma_{000} + \sum_{q=1}^{Q_0} (\gamma_{0qs} * X_{qi}) + [e_{tis} + r_{0is} + u_{00s}] \cdot \tag{8}$$

Due to the removal of time from the model but the inclusion of possible nesting within school district, the subscripts also change slightly. Instead of *y* being the score at time *t* for individual *i* within school *s*, *y* is now the score for individual *i* within school *s* within district *d*. That is, the outcome is now y_{isd}. Given this change in nesting structure, Equation 8 is more appropriately written as:

$$y_{isd} = \gamma_{000} + \sum_{q=1}^{Q_0} (\gamma_{0qd} * X_{qi}) + [e_{isd} + r_{0sd} + u_{00d}] \cdot \tag{9}$$

The interpretation of the growth and difference model parameters generalizes logically to the single time point model parameters. As with the two aforementioned models, for a given outcome, if supported by the data, one or more additional random effects may be constrained to zero.