In the following report, Hanover Research analyzes the demographic data, academic data, as well as the program data and evaluates the effects of the Reading Recovery (RR) program implemented in Metro Nashville Public Schools (MNPS). In the analysis, Hanover focuses on two cohorts of students – those who were in Grade 1 in 2014-15 and 2015-16, respectively.
EXECUTIVE SUMMARY

INTRODUCTION

In this report, Hanover Research (Hanover) evaluates the effects of the Reading Recovery (RR) program implemented in Metro Nashville Public Schools (MNPS). The data used in the analysis consist of student demographic information (e.g., gender, race/ethnicity, English learner status, Special Education status, free or reduced-price lunch status, and immigration status), performance on standardized assessments (i.e., Text Level Assessments, MAP, TN Ready, Fast Bridge), as well as RR program participation data. In our case, we only focus on two cohorts of first grade participants: 1) those who completed the program in 2014-15 and 2) those who completed the program in 2015-16. Hanover further segments each cohort of students by the semester in which they participated in the RR program.

In our analysis, we use the propensity score matching (PSM) technique to construct a control group and compare program participants with the matched similar non-participants on their academic performance in Grade 1 spring semester, as well as the performance in different semesters in later grade levels. The participants and matched non-participants are similar in terms of their demographic characteristics and academic performance in Grade 1.¹

The report comprises two sections:

- **Section I: Data Overview and Methodology** summarizes the data and methodology used in the evaluation.
- **Section II: Comparison between Participants and Matched Non-Participants** presents the comparison between participants and similar non-participants on their post-program academic performance in Grade 1, as well as in the later grade levels. We compare the outcomes between participants and matched non-participants who were identified using the propensity score matching technique.

RECOMMENDATIONS

- MNPS should consider additional follow-up work to determine whether there are improvements that can be made to ensure that more Reading Recovery participants can meet the criteria needed to be discontinued, rather than simply completing the program without being discontinued. We find that discontinued students see much more positive results after exiting the program than non-discontinued students, who continue to lag behind non-participating students.

¹ Please note that we control for fall characteristics and performance for fall semester students, and winter characteristics and performance for spring semester students.
KEY FINDINGS

- Overall, fall 2014-15 RR participants underperformed relative to non-participants across a range of outcomes, but participants in and after the 2014-15 spring semester tended to outperform non-participants in Grade 1, with less consistent results in later grade levels.

- RR participants who were discontinued from the program saw much more positive results than those who completed the program but were not discontinued.
  - Discontinued participants who began RR after fall 2014-15 consistently outperformed non-participants on the TLAs in Grades 1 through 4. There was also some evidence for positive results on Grade 2 CBMR scores and Grade 3 MAP scores, though these results were less consistent.
  - Non-discontinued RR participants consistently underperformed relative to matched non-participants across all assessments and grade levels, with the gap in TLA results widening in later grade levels.
SECTION I: DATA OVERVIEW AND METHODOLOGY

In this report, Hanover Research (Hanover) analyzes the effects of the Reading Recovery (RR) program implemented in MNPS among Grade 1 students in 2014-15 and 2015-16. Student outcomes that we examine include students’ post-program performance on various standardized assessments (TLA, MAP, Fast Bridge, and TN Ready). In this section, we discuss the data and methodology used in the evaluation.

DATA OVERVIEW

The data provided by MNPS comprise of two major components: pre-program data and post-program data. Hanover uses the pre-program data and the propensity score matching (PSM) technique to construct a control group. Then, we utilize the post-program data and compare program participants with the matched similar non-participants on post-program academic performance. The pre-program data provided by MNPS comprise of the following pieces:

- **Program participation data** – include a list of the students who participated in the RR program in either fall or spring semester in 2014-15 and 2015-16. This includes data on whether each student completed the program. Students could complete the RR program either by meeting the program’s criteria for success and being “discontinued” or by participating in the program for 20 weeks without meeting these criteria. Hanover only includes students who completed the program as participants, regardless of whether they were discontinued or not.

- **Assessment data** – include students’ Text Level Assessment results in the fall and spring semesters in 2014-15 and 2015-16. Hanover converted the instructional level to numeric score that ranges from zero to 26.\(^2\) For 10 students with multiple records in the same semester and year, Hanover keeps the record with higher score.

- **Enrollment data** – include students’ enrollment (e.g., grade level, school membership) and demographic (e.g., gender, race/ethnicity) data in the fall and spring semesters in 2014-15 and 2015-16.\(^3\) We focus on Grade 1 students in our study.

- **Demographic data** – include students’ demographic information in the fall and spring semesters in 2014-15 and 2015-16. The datasets are as follow:
  - **Immigrant Migrant data** include students’ immigrant and migrant status.
  - **English language learner (ELL) data** include students’ ELL status. Students are grouped into three categories: ELL, former and/or transitional ELL, and non-ELL (non-English language background).\(^4\)

---

\(^2\) The transformation rule is Pre-A = 0, A = 1, B = 2, C = 3, ……., Z = 26.

\(^3\) When there are multiple datasets of the same type in a semester (e.g., “EOY enrollment 5 August 4 2014.dta” and “EOY enrollment 5 September 5 2014.dta”), Hanover keeps the information in the more recent dataset if any inconsistencies arise. This assumption also applies for Immigrant migrant data, ELL data, special education data, as well as household income data.

\(^4\) Students who were eligible but waived ELL services are categorized as non-ELL students.
Special education (SPED) data include students’ SPED status. Students who cannot be found in the SPED datasets are categorized as non-SPED students.

Household data include students’ household size and income level information. To be eligible for free or reduced-price lunch (FRPL), a student needs to come from a family with an income level falling below 185 percent of the national poverty threshold, or have a valid direct certification date. Students who could not be found in the household data are categorized as non-FRPL (i.e., paid lunch).

The post-program data provided by MNPS comprise of the following components:

- **Text Level Assessment (TLA) data** – include students’ TLA outcomes in the following year-semester combinations: spring in 2014-15, spring in 2015-16, fall, winter, and spring in 2016-17, fall and winter in 2017-18. Hanover converts the instructional level to numeric score that ranges from zero to 26. If a student has multiple records in the same year and semester, Hanover keeps the one with higher instruction level.

- **Measures of Academic Progress (MAP) data** – include students’ MAP outcomes in 2016-17 winter and spring, as well as in 2017-18 fall. In the analysis, Hanover focuses on the MAP reading RIT scores.

- **Fast Bridge Curriculum-Based Measurement for Reading (CBMR) data** – include students’ Fast Bridge CBMR results in 2016-17 fall, winter, and spring. For this analysis, we focus on the district-level percentile as the outcome. For 16 students who take the same test multiple times within the same semester, Hanover keeps the record with the higher score.

- **TN Ready data** – include students’ TN Ready data in Grade 3 in 2016-17. Hanover focuses on the ELA TN Ready test in this analysis.

Figure 1.1 summarizes the academic outcome availability by grade level and year. Hanover uses the pre-program data (i.e., assessment, demographics, program participation) as the anchor file, and merges in post-program data components. Thus, we excluded students who do not have pre-program demographic and assessment data. As discussed previously, we only focus on two cohorts of students – the first grade students who completed the program in 2014-15, spring in 2015-16, fall, winter, and spring in 2016-17, fall and winter in 2017-18. Hanover converts the instructional level to numeric score that ranges from zero to 26. If a student has multiple records in the same year and semester, Hanover keeps the one with higher instruction level.

- In 2015-16, due to a lack of household size and income level information in the fall semester, we use information from the spring semester in that year. In addition, in the spring semester in 2015-16, Hanover capped the household size at 12, and any values larger than 12 are considered as outliers, though all students with these household sizes still met criteria for being classified as receiving free or reduced-price lunch.

- Similar to the pre-program data, the transformation rule is Pre-A = 0, A = 1, B = 2, ......., Z = 26.

- In addition to the CBMR results, the Fast Bridge datasets also include the following two tests: Early Reading and Adaptive Reading. These two tests are excluded due to a relatively small sample size.

- Hanover excludes around 3.4 percent of the total records with an invalid test score (i.e., these students have a value of zero in the "Valid Test" variable).

- The TN Ready datasets also include some records for other tests (i.e., ELA, MSAA, Alternate Science). Hanover excludes these tests for the purpose of this analysis.

---

5 Students with an SPED ending date earlier than the SPED starting date are excluded from the analysis. In addition, in the fall semester, if a student is categorized as SPED between August 7 and September 1, then he/she is marked as SPED in that semester. In spring semester, if a student is categorized as SPED between January 6 and February 1, then he/she is marked as SPED in that semester.

6 In 2015-16, due to a lack of household size and income level information in the fall semester, we use information from the spring semester in that year. In addition, in the spring semester in 2015-16, Hanover capped the household size at 12, and any values larger than 12 are considered as outliers, though all students with these household sizes still met criteria for being classified as receiving free or reduced-price lunch.

7 Similar to the pre-program data, the transformation rule is Pre-A = 0, A = 1, B = 2, ......., Z = 26.

8 In addition to the CBMR results, the Fast Bridge datasets also include the following two tests: Early Reading and Adaptive Reading. These two tests are excluded due to a relatively small sample size.

9 Hanover excludes around 3.4 percent of the total records with an invalid test score (i.e., these students have a value of zero in the “Valid Test” variable).

10 The TN Ready datasets also include some records for other tests (i.e., ELA, MSAA, Alternate Science). Hanover excludes these tests for the purpose of this analysis.
2014-15 and 2015-16. In addition, in this phase of the analysis, we only focus on students who were on-track (i.e., students who have not been retained or skipped grade levels). After the data cleaning steps described above, there are 86, 37, 146, and 110 unique participants in 2014-15 fall, 2014-15 spring, 2015-16 fall, and 2015-16 spring, respectively; and there are 1,720; 1,206; 1,136; and 1,477 unique matched non-participants in 2014-15 fall, 2014-15 spring, 2015-16 fall, and 2015-16 spring, respectively (Figure 1.2). In the spring semester, nearly all RR completers are discontinued, likely because students who do not begin the program until the spring semester mostly do not have time to complete a full 20 weeks in the program before the end of the school year.

Figure 1.1: Outcome Availability by Year, Grade-Level, and Participating Year

<table>
<thead>
<tr>
<th>Year</th>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>Grade 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-15</td>
<td>TLA - Spring</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015-16</td>
<td>TLA - Spring</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016-17</td>
<td>TLA - Fall, Spring, Winter</td>
<td>TLA - Fall, Spring, Winter</td>
<td>MAP - Fall, Winter, Spring</td>
<td>CBMR - Fall, Winter, Spring</td>
</tr>
<tr>
<td>2017-18</td>
<td>TLA - Fall, Winter</td>
<td>TLA - Fall, Winter</td>
<td>MAP - Winter</td>
<td></td>
</tr>
</tbody>
</table>

Green highlighting indicates the year in which the cohort participated in Reading Recovery.

Figure 1.2: Unique Number of Students by Year and Participating Semester in Grade 1

<table>
<thead>
<tr>
<th>Year</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fall</td>
<td>Spring</td>
</tr>
<tr>
<td>2014-15</td>
<td>86</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>Matched Non-Participant</td>
<td>1,720</td>
<td>1,206</td>
</tr>
<tr>
<td>2015-16</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>100</td>
</tr>
<tr>
<td>Matched Non-Participant</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
METHODOLOGY

OUTCOME

As discussed in the Data Overview subsection, in this analysis, Hanover compares the post-program academic performance (from spring in Grade 1 to Grade 4 for Cohort 1 students; from spring in Grade 1 to Grade 3 for Cohort 2 students) between RR participants and matched non-participants. The outcomes of interest are listed in Figure 1.1.

COMPARING THE PARTICIPANTS AND NON-PARTICIPANTS

For each year and semester, we conduct the PSM separately. A critical step in evaluating the effects of participating in the Reading Recovery (RR) program on student outcomes is to identify an appropriate comparison group with which the participants’ outcomes can be compared. Students were selected for RR eligibility based on prior measures of reading proficiency, which will influence students’ post-program performance. In addition, other student characteristics may be associated with the likelihood of RR participation, and these characteristics may also be related to the performance outcomes. Thus, a direct comparison between participants and non-participants may not reflect the true program effect, but might instead measure pre-existing differences between the two groups.

Please note that for students who participated in the program in different semesters (i.e., fall-semester participants and spring-semester participants), we utilize different pre-program data to compare between participants and non-participants. We use the demographic and assessment data in fall for fall participants; and the demographic and assessment data in winter for spring participants. 11 Figure 1.3 to Figure 1.6 display students’ pre-program characteristics in Grade 1 by program participation status (i.e., fall characteristics and performance for semester one students; winter characteristics and performance for semester two students). T-tests are conducted to examine the statistical significance in the mean differences. As the figures below show, the participants and non-participants differ in many aspects.

PROPENSITY SCORE MATCHING (PSM)

We employ a propensity score matching technique to construct a comparison group referred to as the “matched non-participants.” This technique allows us to construct a group of non-participants who shared similar observable characteristics to the participants in Grade 1.

---

11 We utilize winter data to compare between spring participants and non-participants because those are the most recent data prior to the program. The correlation between fall and winter assessment results is high but not overwhelmingly strong among Grade 1 students in 2014-15 (0.667). Thus, an alternative method would be to control for both fall and winter characteristics and performance for spring participants, especially for Grade 1 students in 2014-15. However, out of the 37 program participants in the spring semester in 2014-15, 11 do not have fall TLA assessment outcomes. In order to include those students and maximize the number of participants included in the analysis, we ultimately did not control for fall assessment performance in the PSM or in the regression models on the post-program outcome performance.
Implementing the propensity score matching method involves estimating a logistic regression model that predicts whether a student participated in the RR program, using demographic characteristics (including gender, race/ethnicity, English proficiency level, free or reduced-price lunch status, special education status, immigrant status, and age) and baseline TLA assessment performance (numeric instructional level) in Grade 1 as predictor variables. As discussed previously, we control for fall characteristics and performance for semester one students, and winter characteristics and performance for semester two students. The resulting equation assigns each student a propensity score that represents his or her predicted probability of participating in the RR program based on these characteristics. Essentially, this number represents our best estimate of how likely each student is to participate in Reading Recovery, based only on the information that is available prior to the start of participation.

Each participating student is then matched with non-participating students with similar propensity scores. While many studies match each participant with a specific number of non-participants (frequently using one-to-one matching), recent research has found that this matching technique can lead to poor results because it frequently leads to the exclusion of non-participants who are similar enough to participants to be valuable for comparison purposes, essentially throwing out data for some non-participants at random. To mitigate this issue, we instead use radius matching to include all non-participants whose propensity score falls within a specific range of each participant. By including all non-participants within this radius rather than matching only one or a handful of the non-participants, this radius matching technique helps to mitigate some of the potential issues with nearest neighbor matching by ensuring that high-quality matches are not excluded from the analysis due to artificial limitations imposed by the matching process. We set a 0.0001 radius, which indicates that each participant is matched with the non-participants whose propensity score is within 0.0001 of the participant’s score (in either direction).

We conduct the matching for each semester and year separately. In addition, 15 participants who had missing fall benchmark scores are excluded from the matching and subsequent analysis.

We note that the propensity score matching technique has its limitations. PSM does not take any unobservable factors into account. Thus, even though the matched non-participants are similar to the participants on all observed variables, they may still differ from the non-

---

14 Hanover sets the radius based on two considerations: 1) if the radius is too wide, we may end up including some non-participants who are not especially similar to the participants in terms of pre-program demographic characteristics and assessment performance; 2) if the radius is too narrow, some participants may not be matched to any non-participants and would be excluded from our analysis. Ultimately, the radius is set at 0.0001 to balance these two factors. This results in only 22 participants being excluded due to an inability to find a matching non-participant.
participants in other unmeasured aspects that are related to participants and student outcomes.

**Regression after the PSM**

After the matching, participants and non-participants could still be significantly different in terms of the observed characteristics on average. While the matching ensures that each matched non-participant is similar to at least one participant, there may still be group-level differences since some participants will match to many more non-participants than others. After running the PSM, we conduct descriptive analysis to examine and compare the Grade 1 pre-program demographic characteristics and TLA performance between participants and matched non-participants. Figure 1.7 to Figure 1.10 present the descriptive analysis results for participants and matched non-participants. The results indicate that the students are still quite different from each other in terms of demographic characteristics and TLA assessment performance on average. Because of these remaining differences between the two groups on average, we use regression models to compare the outcomes between the treatment and control group based on the constructed sample, while controlling for Grade 1 pre-program demographic characteristics and TLA performance.

For each of the numeric academic outcomes (Figure 1.1), we build linear regression models based on the constructed sample for each cohort of students by participating semester. The explanatory variable of interest is whether a student participated in the program, while controlling for the Grade 1 pre-program demographic characteristics (i.e., gender, race/ethnicity, ELL status, SPED status, immigration status, and FRPL status), as well as pre-program TLA assessment performance. By interpreting the coefficient of the program participation variable, we are able to compare participants and matched non-participants (who are similar in terms of Grade 1 demographic characteristics and TLA assessment performance to some extent) and analyze whether they have significantly different outcomes.

We also fit a second set of models for each outcome, segmenting the program participants by discontinued vs. non-discontinued status. While the models focusing on overall program participation provide the most direct evaluation of the program’s effect on all participating students, these additional models may provide insight into whether outcomes differ for students who meet the program’s success criteria.

---

15 Please note that these tables do not weight observations based on the number of non-participants matched to each participant, which is why students are still significantly different in terms of certain observed characteristics. While each matched non-participant is similar to at least one participant, some participants have many more matched non-participants than others. In the program evaluation analyses to be conducted, we control for characteristics for which differences still exist between the participants and matched non-participants.
### Descriptive Analysis Results before the Propensity Score Matching

**Figure 1.3: Student Characteristics before the PSM – Participants vs. Non-Participants**

#### 2014-15 Semester One

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
<th>DIFFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
<td>COUNT</td>
</tr>
<tr>
<td><strong>TLA Outcome</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>91</td>
<td>2.714</td>
<td>6,316</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>104</td>
<td>38.5%</td>
<td>7,897</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>104</td>
<td>0%</td>
<td>7,897</td>
</tr>
<tr>
<td>Asian</td>
<td>104</td>
<td>2.9%</td>
<td>7,897</td>
</tr>
<tr>
<td>Black</td>
<td>104</td>
<td>22.1%</td>
<td>7,897</td>
</tr>
<tr>
<td>Hispanic</td>
<td>104</td>
<td>52.9%</td>
<td>7,897</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>104</td>
<td>0%</td>
<td>7,897</td>
</tr>
<tr>
<td>White</td>
<td>104</td>
<td>22.1%</td>
<td>7,897</td>
</tr>
<tr>
<td>Age</td>
<td>104</td>
<td>6.373</td>
<td>7,897</td>
</tr>
<tr>
<td>ELL</td>
<td>104</td>
<td>59.6%</td>
<td>7,897</td>
</tr>
<tr>
<td>Former ELL</td>
<td>104</td>
<td>0%</td>
<td>7,897</td>
</tr>
<tr>
<td>Special Education</td>
<td>104</td>
<td>6.7%</td>
<td>7,897</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>104</td>
<td>93.3%</td>
<td>7,897</td>
</tr>
<tr>
<td>Immigrant</td>
<td>104</td>
<td>10.6%</td>
<td>7,897</td>
</tr>
</tbody>
</table>

**Figure 1.4: Student Characteristics before the PSM – Participants vs. Non-Participants**

#### 2014-15 Semester Two

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
<th>DIFFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
<td>COUNT</td>
</tr>
<tr>
<td><strong>TLA Outcome</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>37</td>
<td>3.541</td>
<td>7,108</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
<td>51.4%</td>
<td>7,833</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>37</td>
<td>0%</td>
<td>7,833</td>
</tr>
<tr>
<td>Asian</td>
<td>37</td>
<td>2.7%</td>
<td>7,833</td>
</tr>
<tr>
<td>Black</td>
<td>37</td>
<td>27%</td>
<td>7,833</td>
</tr>
<tr>
<td>Hispanic</td>
<td>37</td>
<td>40.5%</td>
<td>7,833</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>37</td>
<td>0%</td>
<td>7,833</td>
</tr>
<tr>
<td>White</td>
<td>37</td>
<td>29.7%</td>
<td>7,833</td>
</tr>
<tr>
<td>Age</td>
<td>37</td>
<td>6.398</td>
<td>7,833</td>
</tr>
<tr>
<td>ELL</td>
<td>37</td>
<td>40.5%</td>
<td>7,833</td>
</tr>
<tr>
<td>Former ELL</td>
<td>37</td>
<td>0%</td>
<td>7,833</td>
</tr>
<tr>
<td>Special Education</td>
<td>37</td>
<td>2.7%</td>
<td>7,833</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>37</td>
<td>100%</td>
<td>7,833</td>
</tr>
<tr>
<td>Immigrant</td>
<td>37</td>
<td>8.1%</td>
<td>7,833</td>
</tr>
</tbody>
</table>
Figure 1.5: Student Characteristics before the PSM – Participants vs. Non-Participants

2015-16 Semester One

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
<th>DIFFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
<td>COUNT</td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>167</td>
<td>1.162</td>
<td>6,234</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>167</td>
<td>43.1%</td>
<td>7,066</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>167</td>
<td>0%</td>
<td>7,066</td>
</tr>
<tr>
<td>Asian</td>
<td>167</td>
<td>3.6%</td>
<td>7,066</td>
</tr>
<tr>
<td>Black</td>
<td>167</td>
<td>36.5%</td>
<td>7,066</td>
</tr>
<tr>
<td>Hispanic</td>
<td>167</td>
<td>40.7%</td>
<td>7,066</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>167</td>
<td>0%</td>
<td>7,066</td>
</tr>
<tr>
<td>White</td>
<td>167</td>
<td>19.2%</td>
<td>7,066</td>
</tr>
<tr>
<td>Age</td>
<td>167</td>
<td>6.413</td>
<td>7,066</td>
</tr>
<tr>
<td>ELL</td>
<td>167</td>
<td>41.3%</td>
<td>7,066</td>
</tr>
<tr>
<td>Former ELL</td>
<td>167</td>
<td>0%</td>
<td>7,066</td>
</tr>
<tr>
<td>Special Education</td>
<td>167</td>
<td>4.2%</td>
<td>7,066</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>167</td>
<td>91%</td>
<td>7,066</td>
</tr>
<tr>
<td>Immigrant</td>
<td>167</td>
<td>6.6%</td>
<td>7,066</td>
</tr>
</tbody>
</table>

Figure 1.6: Student Characteristics before the PSM – Participants vs. Non-Participants

2015-16 Semester Two

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
<th>DIFFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
<td>COUNT</td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>110</td>
<td>4.309</td>
<td>6,337</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>112</td>
<td>47.3%</td>
<td>7,106</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>112</td>
<td>0%</td>
<td>7,106</td>
</tr>
<tr>
<td>Asian</td>
<td>112</td>
<td>1.8%</td>
<td>7,106</td>
</tr>
<tr>
<td>Black</td>
<td>112</td>
<td>38.4%</td>
<td>7,106</td>
</tr>
<tr>
<td>Hispanic</td>
<td>112</td>
<td>40.2%</td>
<td>7,106</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>112</td>
<td>0%</td>
<td>7,106</td>
</tr>
<tr>
<td>White</td>
<td>112</td>
<td>19.6%</td>
<td>7,106</td>
</tr>
<tr>
<td>Age</td>
<td>112</td>
<td>6.447</td>
<td>7,106</td>
</tr>
<tr>
<td>ELL</td>
<td>112</td>
<td>45.5%</td>
<td>7,106</td>
</tr>
<tr>
<td>Former ELL</td>
<td>112</td>
<td>0%</td>
<td>7,106</td>
</tr>
<tr>
<td>Special Education</td>
<td>112</td>
<td>1.8%</td>
<td>7,106</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>112</td>
<td>92%</td>
<td>7,106</td>
</tr>
<tr>
<td>Immigrant</td>
<td>112</td>
<td>7.1%</td>
<td>7,106</td>
</tr>
</tbody>
</table>
DESCRIPTIVE ANALYSIS RESULTS AFTER THE PROPENSITY SCORE MATCHING

Figure 1.7: Student Characteristics after the PSM – Participants vs. Non-Participants
2014-15 Semester One

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
</tr>
<tr>
<td>TLA Outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>87</td>
<td>2.805</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>87</td>
<td>39.1%</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>87</td>
<td>0%</td>
</tr>
<tr>
<td>Asian</td>
<td>87</td>
<td>3.4%</td>
</tr>
<tr>
<td>Black</td>
<td>87</td>
<td>25.3%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>87</td>
<td>50.6%</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>87</td>
<td>0%</td>
</tr>
<tr>
<td>White</td>
<td>87</td>
<td>20.7%</td>
</tr>
<tr>
<td>Age</td>
<td>87</td>
<td>6.397</td>
</tr>
<tr>
<td>ELL</td>
<td>87</td>
<td>56.3%</td>
</tr>
<tr>
<td>Former ELL</td>
<td>87</td>
<td>0%</td>
</tr>
<tr>
<td>Special Education</td>
<td>87</td>
<td>8%</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>87</td>
<td>93.1%</td>
</tr>
<tr>
<td>Immigrant</td>
<td>87</td>
<td>8%</td>
</tr>
</tbody>
</table>

Figure 1.8: Student Characteristics after the PSM – Participants vs. Non-Participants
2014-15 Semester Two

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
</tr>
<tr>
<td>TLA Outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>37</td>
<td>3.541</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
<td>51.4%</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>37</td>
<td>0%</td>
</tr>
<tr>
<td>Asian</td>
<td>37</td>
<td>2.7%</td>
</tr>
<tr>
<td>Black</td>
<td>37</td>
<td>27%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>37</td>
<td>40.5%</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>37</td>
<td>0%</td>
</tr>
<tr>
<td>White</td>
<td>37</td>
<td>29.7%</td>
</tr>
<tr>
<td>Age</td>
<td>37</td>
<td>6.398</td>
</tr>
<tr>
<td>ELL</td>
<td>37</td>
<td>40.5%</td>
</tr>
<tr>
<td>Former ELL</td>
<td>37</td>
<td>0%</td>
</tr>
<tr>
<td>Special Education</td>
<td>37</td>
<td>2.7%</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>37</td>
<td>100%</td>
</tr>
<tr>
<td>Immigrant</td>
<td>37</td>
<td>8.1%</td>
</tr>
</tbody>
</table>
Figure 1.9: Student Characteristics after the PSM – Participants vs. Non-Participants

2015-16 Semester One

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
</tr>
<tr>
<td>TLA Outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>149</td>
<td>1.268</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>149</td>
<td>42.3%</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>149</td>
<td>0%</td>
</tr>
<tr>
<td>Asian</td>
<td>149</td>
<td>2%</td>
</tr>
<tr>
<td>Black</td>
<td>149</td>
<td>38.3%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>149</td>
<td>40.3%</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>149</td>
<td>0%</td>
</tr>
<tr>
<td>White</td>
<td>149</td>
<td>19.5%</td>
</tr>
<tr>
<td>Age</td>
<td>149</td>
<td>6.408</td>
</tr>
<tr>
<td>ELL</td>
<td>149</td>
<td>43%</td>
</tr>
<tr>
<td>Former ELL</td>
<td>149</td>
<td>0%</td>
</tr>
<tr>
<td>Special Education</td>
<td>149</td>
<td>4.7%</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>149</td>
<td>90.6%</td>
</tr>
<tr>
<td>Immigrant</td>
<td>149</td>
<td>6.7%</td>
</tr>
</tbody>
</table>

Figure 1.10: Student Characteristics after the PSM – Participants vs. Non-Participants

2015-16 Semester Two

<table>
<thead>
<tr>
<th>CHARACTERISTICS</th>
<th>PARTICIPANTS</th>
<th>NON-PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>COUNT</td>
<td>MEAN</td>
</tr>
<tr>
<td>TLA Outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeric Instructional Level</td>
<td>110</td>
<td>4.309</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>110</td>
<td>48.2%</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>110</td>
<td>0%</td>
</tr>
<tr>
<td>Asian</td>
<td>110</td>
<td>1.8%</td>
</tr>
<tr>
<td>Black</td>
<td>110</td>
<td>38.2%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>110</td>
<td>40%</td>
</tr>
<tr>
<td>Hawaiian or Other Pacific Islander</td>
<td>110</td>
<td>0%</td>
</tr>
<tr>
<td>White</td>
<td>110</td>
<td>20%</td>
</tr>
<tr>
<td>Age</td>
<td>110</td>
<td>6.448</td>
</tr>
<tr>
<td>ELL</td>
<td>110</td>
<td>45.5%</td>
</tr>
<tr>
<td>Former ELL</td>
<td>110</td>
<td>0%</td>
</tr>
<tr>
<td>Special Education</td>
<td>110</td>
<td>1.8%</td>
</tr>
<tr>
<td>Free or Reduced-Price Lunch</td>
<td>110</td>
<td>92.7%</td>
</tr>
<tr>
<td>Immigrant</td>
<td>110</td>
<td>7.3%</td>
</tr>
</tbody>
</table>
SECTION II: COMPARISON BETWEEN PARTICIPANTS AND MATCHED NON-PARTICIPANTS

In this section, we present and discuss the regression results that compare the post-program academic performance between RR participants and matched non-participants who were identified using the propensity score matching method.

To construct a comparison group, we employ the propensity score matching technique. However, Hanover observed that the participants and matched non-participants are still significantly different in terms of pre-program demographic characteristics and TLA performance in Grade 1. Thus, Hanover builds linear regression models to compare the academic outcome variables of interest between the treatment and control group based on the constructed sample, while controlling for pre-program demographic characteristics (e.g., gender, race/ethnicity, ELL status, SPED status, FRPL status, immigration status) and TLA performance in Grade 1. For fall participants, we control for their fall characteristics and performance; while for spring participants, we control for their winter characteristics and performance.

In this section, we only present the coefficients of the explanatory variable of interest (i.e., program participation status) and do not present the coefficients for the control variables. Please refer to the accompanying data supplement for the complete regression result tables.

MAIN TAKEAWAYS

- Overall, fall 2014-15 RR participants underperformed relative to non-participants across a range of outcomes, but participants in and after the 2014-15 spring semester tended to outperform non-participants in Grade 1, with less consistent results in later grade levels.
  - Fall participants in 2015-16 outperform matched non-participants at the end of Grade 1 by about one instructional level, on average, and in the fall of Grade 2, by about half an instructional level, but differences in later grade levels are mostly not significant.
  - Overall results are more positive for spring participants, but this is likely driven by the fact that nearly all spring RR participants who complete the program are discontinued, since students who only begin the program in the spring generally do not have time to complete the program without being discontinued before the end of the school year.
- **RR participants who were discontinued from the program saw much more positive results than those who completed the program but were not discontinued.**
  - Discontinued participants consistently outperformed matched non-participants in the spring of Grade 1, with TLA results roughly one to two levels higher than matched non-participants.
  - In Grade 2, discontinued RR participants in terms after fall 2014-15 had higher TLA results than non-participants. There was also less consistent evidence for a positive effect on Grade 2 CBMR scores, though this outcome was only available for 2015-16 participants and differences were not statistically significant for spring 2015-16 participants on the winter and spring assessments. There were no significant differences in Grade 2 MAP RIT scores.
  - In Grades 3 and 4, discontinued participants consistently outperformed participants on the TLA. Differences on the MAP and CBMR were not consistently significant, though discontinued 2015-16 participants did have higher fall MAP reading scores in Grade 3.

- **Non-discontinued RR participants consistently underperformed relative to matched non-participants across all assessments and grade levels.**
  - The gap between non-discontinued RR completers and non-participants on the TLA widens with each successive grade level. For fall 2014-15 participants, for example, there is an average gap of roughly 2.5 instructional levels on the spring TLA in Grades 1 and 2, which widens to 3.3 in spring of Grade 3 and 3.9 by winter of Grade 4.
**FIGURES**

**Figure 2.1: Regression Analysis Results – 2014-15 Fall Participants (Part I)**

<table>
<thead>
<tr>
<th></th>
<th>2014-15 Grade 1</th>
<th>2015-16 Grade 2</th>
<th>2016-17 Grade 3</th>
<th>MAP READING RIT Score Winter</th>
<th>MAP READING RIT Score Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program Participation (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>-1.5092***</td>
<td>-1.7158***</td>
<td>-1.8794***</td>
<td>-2.1060***</td>
<td>-2.2508***</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,806</td>
<td>1,621</td>
<td>1,503</td>
<td>1,486</td>
<td>1,471</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4372</td>
<td>0.3406</td>
<td>0.3416</td>
<td>0.3219</td>
<td>0.2982</td>
</tr>
<tr>
<td><strong>Program Participation and Discontinued Status (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued</td>
<td>0.8856*</td>
<td>0.2316</td>
<td>0.2532</td>
<td>0.1286</td>
<td>0.2366</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,806</td>
<td>1,621</td>
<td>1,503</td>
<td>1,486</td>
<td>1,471</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4448</td>
<td>0.3447</td>
<td>0.3457</td>
<td>0.3269</td>
<td>0.3040</td>
</tr>
</tbody>
</table>

Note: * p < 0.10, ** p < 0.05, *** p < 0.01.
**Figure 2.2: Regression Analysis Results – 2014-15 Fall Participants (Part II)**

<table>
<thead>
<tr>
<th></th>
<th>2016-17 Grade 3</th>
<th></th>
<th></th>
<th>2017-18 Grade 4</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBMR DISTRICT PERCENTILE FALL</td>
<td>CBMR DISTRICT PERCENTILE WINTER</td>
<td>CBMR DISTRICT PERCENTILE SPRING</td>
<td>TN READY ELA SCALED SCORE</td>
<td>TLA FALL</td>
<td>TLA WINTER</td>
</tr>
<tr>
<td><strong>Program Participation (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,488</td>
<td>1,481</td>
<td>1,468</td>
<td>1,456</td>
<td>1,363</td>
<td>1,358</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.3055</td>
<td>0.3031</td>
<td>0.2869</td>
<td>0.2982</td>
<td>0.3103</td>
<td>0.2918</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Program Participation and Discontinued Status (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued Participants</td>
<td>0.6188</td>
<td>-0.5064</td>
<td>-0.4081</td>
<td>-3.9073</td>
<td>0.3791</td>
<td>0.3075</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,488</td>
<td>1,481</td>
<td>1,468</td>
<td>1,456</td>
<td>1,363</td>
<td>1,358</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.3121</td>
<td>0.3097</td>
<td>0.2917</td>
<td>0.3015</td>
<td>0.3176</td>
<td>0.3000</td>
</tr>
</tbody>
</table>

Note: * p < 0.10, ** p < 0.05, *** p < 0.01.
### Figure 2.3: Regression Analysis Results – 2014-15 Spring Participants (Part I)

<table>
<thead>
<tr>
<th></th>
<th>2014-15 GRADE 1</th>
<th>2015-16 GRADE 2</th>
<th>2016-17 GRADE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLA SPRING</td>
<td>1.8344***</td>
<td>1.6819***</td>
<td>1.6143***</td>
</tr>
<tr>
<td>TLA SPRING</td>
<td>1.6143***</td>
<td>1.0236**</td>
<td>1.1043**</td>
</tr>
<tr>
<td>TLA FALL</td>
<td>1.0236**</td>
<td>1.1043**</td>
<td>1.0313</td>
</tr>
<tr>
<td>TLA WINTER</td>
<td>1.0313</td>
<td>1.0313</td>
<td>-0.6038</td>
</tr>
<tr>
<td>MAP READING</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIT SCORE SPRING</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WINTER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.7040</td>
<td>0.5380</td>
<td>0.4973</td>
</tr>
<tr>
<td></td>
<td>0.5233</td>
<td>0.4973</td>
<td>0.5068</td>
</tr>
<tr>
<td></td>
<td>1.006</td>
<td>0.4589</td>
<td>0.4401</td>
</tr>
</tbody>
</table>

Program Participation (Ref: Matched Non-Participants)

- **Participants:** 1,243
- **Variables Controlled:** Pre-program winter TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership
- **Observations:** 1,032
- **R-squared:** 0.7040

Note: * p < 0.10, ** p < 0.05, *** p < 0.01. Because all Spring 2014-15 participants who completed the program were discontinued, comparisons between discontinued and non-discontinued students are not possible for this semester.

### Figure 2.4: Regression Analysis Results – 2014-15 Spring Participants (Part II)

<table>
<thead>
<tr>
<th></th>
<th>2016-17 GRADE 3</th>
<th>2017-18 GRADE 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLA SPRING</td>
<td>3.7296</td>
<td>3.5862</td>
</tr>
<tr>
<td>TLA SPRING</td>
<td>2.5917</td>
<td>1.1452</td>
</tr>
<tr>
<td>TLA SPRING</td>
<td>1.2778**</td>
<td>1.9683***</td>
</tr>
<tr>
<td>TLA SPRING</td>
<td>1.013</td>
<td>0.3716</td>
</tr>
<tr>
<td>MAP READING</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIT SCORE SPRING</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WINTER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4677</td>
<td>0.4497</td>
</tr>
<tr>
<td></td>
<td>0.4129</td>
<td>0.4020</td>
</tr>
<tr>
<td></td>
<td>0.4791</td>
<td>0.4955</td>
</tr>
<tr>
<td></td>
<td>0.3716</td>
<td>0.4373</td>
</tr>
</tbody>
</table>

Program Participation (Ref: Matched Non-Participants)

- **Participants:** 1,036
- **Variables Controlled:** Pre-program winter TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership
- **Observations:** 951
- **R-squared:** 0.4677

Note: * p < 0.10, ** p < 0.05, *** p < 0.01. Because all Spring 2014-15 participants who completed the program were discontinued, comparisons between discontinued and non-discontinued students are not possible for this semester.
### Figure 2.5: Regression Analysis Results – 2015-16 Fall Participants (Part I)

<table>
<thead>
<tr>
<th></th>
<th>2015-16 Grade 1</th>
<th>2016-17 Grade 2</th>
<th>MAP READING RIT Score Winter</th>
<th>MAP READING RIT Score Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLA Spring</td>
<td>0.9779***</td>
<td>0.5161*</td>
<td>0.1592</td>
<td>0.4314</td>
</tr>
<tr>
<td>TLA Fall</td>
<td>0.5161*</td>
<td>1.158</td>
<td>1.133</td>
<td>1.122</td>
</tr>
<tr>
<td>TLA Winter</td>
<td>0.1592</td>
<td>1.133</td>
<td>1.122</td>
<td>1.092</td>
</tr>
<tr>
<td>TLA Spring</td>
<td>0.4314</td>
<td>1.122</td>
<td>1.092</td>
<td>1.068</td>
</tr>
<tr>
<td>Program Participation (Ref: Matched Non-Participants)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>0.9779***</td>
<td>0.5161*</td>
<td>0.1592</td>
<td>0.4314</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,282</td>
<td>1,186</td>
<td>1,158</td>
<td>1,133</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4758</td>
<td>0.4536</td>
<td>0.4443</td>
<td>0.4363</td>
</tr>
<tr>
<td>Program Participation and Discontinued Status (Ref: Matched Non-Participants)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued Participants</td>
<td>2.3261***</td>
<td>2.1355***</td>
<td>1.9084***</td>
<td>2.2082***</td>
</tr>
<tr>
<td>Non-Discontinued Participants</td>
<td>-0.4045</td>
<td>-1.0584***</td>
<td>-1.6836***</td>
<td>-1.4667***</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,282</td>
<td>1,186</td>
<td>1,158</td>
<td>1,133</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4944</td>
<td>0.4742</td>
<td>0.4699</td>
<td>0.4604</td>
</tr>
</tbody>
</table>

Note: * p < 0.10, ** p < 0.05, *** p < 0.01.
### Figure 2.6: Regression Analysis Results – 2015-16 Fall Participants (Part II)

<table>
<thead>
<tr>
<th></th>
<th>2016-17 Grade 2</th>
<th></th>
<th></th>
<th>2017-18 Grade 3</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBMR District Percentile Fall</td>
<td>CBMR District Percentile Winter</td>
<td>CBMR District Percentile Spring</td>
<td>TLA Fall</td>
<td>TLA Winter</td>
<td>MAP Reading RIT Score Fall</td>
<td></td>
</tr>
<tr>
<td><strong>Variables Controlled</strong></td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>-0.5608</td>
<td>-2.1909</td>
<td>-1.2491</td>
<td>-0.2121</td>
<td>-0.1609</td>
<td>-1.1493</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,161</td>
<td>1,142</td>
<td>1,135</td>
<td>1,049</td>
<td>1,045</td>
<td>1,041</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4601</td>
<td>0.4287</td>
<td>0.4167</td>
<td>0.4166</td>
<td>0.4054</td>
<td>0.4094</td>
<td></td>
</tr>
<tr>
<td><strong>Variables Controlled</strong></td>
<td>Pre-program fall TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued Participants</td>
<td>7.6086***</td>
<td>7.5733***</td>
<td>8.1256***</td>
<td>1.3352***</td>
<td>1.4297***</td>
<td>2.9624*</td>
<td></td>
</tr>
<tr>
<td>Non-Discontinued Participants</td>
<td>-8.6795***</td>
<td>-12.3861***</td>
<td>-11.1870***</td>
<td>-1.9002***</td>
<td>-1.7516***</td>
<td>-5.6094***</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,161</td>
<td>1,142</td>
<td>1,135</td>
<td>1,049</td>
<td>1,045</td>
<td>1,041</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4725</td>
<td>0.4462</td>
<td>0.4325</td>
<td>0.4337</td>
<td>0.4221</td>
<td>0.4173</td>
<td></td>
</tr>
</tbody>
</table>

Note: * p < 0.10, ** p < 0.05, *** p < 0.01.
**Figure 2.7: Regression Analysis Results – 2015-16 Spring Participants (Part I)**

<table>
<thead>
<tr>
<th></th>
<th>2015-16 Grade 1</th>
<th>2016-17 Grade 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TLA SPRING</td>
<td>TLA FALL</td>
</tr>
<tr>
<td><strong>Program Participation (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>1.7522***</td>
<td>1.1448***</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program winter TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,587</td>
<td>1,465</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.6827</td>
<td>0.6212</td>
</tr>
<tr>
<td><strong>Program Participation and Discontinued Status (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued Participants</td>
<td>2.0191***</td>
<td>1.5111***</td>
</tr>
<tr>
<td>Non-Discontinued Participants</td>
<td>-0.9411**</td>
<td>-2.1514***</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program winter TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,587</td>
<td>1,465</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.6880</td>
<td>0.6281</td>
</tr>
</tbody>
</table>

Note: * p < 0.10, ** p < 0.05, *** p < 0.01.
Figure 2.8: Regression Analysis Results – 2015-16 Spring Participants (Part II)

<table>
<thead>
<tr>
<th></th>
<th>2016-17 GRADE 2</th>
<th>2017-18 GRADE 3</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CBMR DISTRICT</td>
<td>CBMR DISTRICT</td>
<td>CBMR DISTRICT</td>
<td>TLA FALL</td>
<td>TLA WINTER</td>
<td>MAP READING</td>
</tr>
<tr>
<td></td>
<td>PERCENTILE FALL</td>
<td>PERCENTILE WINTER</td>
<td>PERCENTILE SPRING</td>
<td></td>
<td></td>
<td>RIT SCORE FALL</td>
</tr>
<tr>
<td><strong>Program Participation (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>1.9008</td>
<td>1.3247</td>
<td>0.2112</td>
<td>0.4616</td>
<td>0.1834</td>
<td>1.9136</td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program winter TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,444</td>
<td>1,422</td>
<td>1,419</td>
<td>1,313</td>
<td>1,322</td>
<td>1,311</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.5767</td>
<td>0.5052</td>
<td>0.4788</td>
<td>0.5088</td>
<td>0.4902</td>
<td>0.4619</td>
</tr>
<tr>
<td><strong>Program Participation and Discontinued Status (Ref: Matched Non-Participants)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discontinued</td>
<td>3.3068**</td>
<td>2.8491</td>
<td>1.7364</td>
<td>0.7314**</td>
<td>0.4264</td>
<td>2.6525*</td>
</tr>
<tr>
<td>Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variables Controlled</td>
<td>Pre-program winter TLA performance, demographic characteristics (gender, race/ethnicity, SPED, ELL, FRPL, Immigration status), and school membership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,444</td>
<td>1,422</td>
<td>1,419</td>
<td>1,313</td>
<td>1,322</td>
<td>1,311</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.5791</td>
<td>0.5086</td>
<td>0.4818</td>
<td>0.5127</td>
<td>0.4939</td>
<td>0.4637</td>
</tr>
</tbody>
</table>

Note: * p < 0.10, ** p < 0.05, *** p < 0.01.
PROJECT EVALUATION FORM

Hanover Research is committed to providing a work product that meets or exceeds client expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.


CAVEAT

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties that extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every client. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Clients requiring such services are advised to consult an appropriate professional.