# California Growth Study

Completed by Educational Testing Service   
for the California Department of Education   
under CAASPP Contract #CN150012

# Executive Summary

This memorandum summarizes an analysis, conducted by Educational Testing Service (ETS) at the request of the California Department of Education (CDE), intended to assist California in its selection of an academic growth measure for inclusion in its accountability system for schools and local educational agencies (LEAs). The CDE, Technical Design Group, and the Technical Advisory Group selected three possible growth models for analysis on the basis of the selection criteria adopted by the State Board of Education. The analysis evaluates the statistical properties of these three candidate student growth measures, which are aggregated to the school and LEA levels. The analysis is not intended to evaluate growth measures for reporting at the individual student level.

The analysis uses data from more than two million California students in grades four through eight during the 2015–16 school year. The data for each student include:

* California Assessment of Student Performance and Progress (CAASSP) Smarter Balanced English language arts/literacy (ELA) and mathematics test scores from the spring 2016 CAASPP administration;
* the corresponding test scores from those same students from the prior year’s CAASPP administration (spring 2015);
* student demographic and program participation characteristics:

**Gender**

Female

Male

**Race/Ethnicity**

American Indian or Alaska Native

Asian

Native Hawaiian or Other Pacific Islander

Filipino

Hispanic or Latino

Black or African American

White

Two or more races

Missing (Unknown race/ethnicity)

**English Learner (EL) Status[[1]](#footnote-2)**

Ever classified EL (Ever EL)

Currently classified EL (Current EL)

Formerly but not currently classified EL (Former EL)

Never classified EL (Never EL)

**Disability Status**

Student with disability

Student with no disability

**Economically Disadvantaged Status[[2]](#footnote-3)**

Socioeconomically disadvantaged

Not socioeconomically disadvantaged

**Homeless Status**

Homeless

Not homeless

**Foster Care Status**

Foster youth

Not foster youth

* the link of each student to his or her school and LEA.

The analysis uses these data to compute, and then compare and contrast, three candidate growth measures aggregated to the school and LEA levels. The goal of the analysis is to use the statistical properties of the measures, along with statistical theory, to identify advantages and disadvantages of the three candidate growth measures for use in California’s accountability system. In addition to these statistical considerations, there may also be practical, political, and logistical considerations that factor into the final selection of a growth measure, but they are beyond the scope of this study.

Each of the three growth measures uses the test score data from both the 2014–15 CAASPP administration (prior-year scores) and 2015–16 CAASPP administration (current-year scores) to compute a measure of achievement progress for each student. For each of the following three measures, the school and LEA aggregate measures are obtained by simply averaging the associated student-level measure:

1. “Change in distance-to-met” (CDTM) answers the question of whether the student is improving relative to proficiency across years and determines if a student is scoring higher relative to the proficiency threshold in the current year than in the previous year. CDTM first computes the “distance-to-met” for the current school year, which equals the student’s current-year test score minus the proficiency threshold for the student’s grade level in the current year. Next, it computes the distance-to-met for the prior school year, which is equal to the student’s test score from the prior school year minus the proficiency threshold for the student’s grade level in the prior year. Finally, it subtracts the prior-year distance-to-met from the current year distance-to-met.
2. “Conditional percentile rank of the gain” (CPR) for a student is defined as the percentile rank corresponding to his/her gain score. It first groups together all students in a grade level with the same prior year score and then computes the gain score (current year score minus prior year score) for each of those students. The gain scores within each group are rank ordered and assigned a percentile rank (rank divided by number of students in the group).
3. “Residual gain” (RG) is the difference between a student’s current test score and a prediction of that score based on the ELA and mathematics test scores from the prior year.

Three criteria were used to evaluate the three candidate growth measures:

1. strength of relationships between each measure and student demographic characteristics,
2. sensitivity of each measure to school grade configuration and assessment content area (mathematics and ELA), and
3. statistical precision of each measure.

Among the three candidate measures, RG has the most desirable properties overall. This conclusion is based on the fact that CPR appears to be more influenced by student demographic characteristics than the other measures, while CDTM has considerably greater sensitivity to both grade configuration and content area than the other measures, as well as slightly lower statistical precision. The rest of the memorandum includes the details and arguments that support these conclusions.

# Growth Study Report

## Purpose

To assist California in its selection of an academic growth component for inclusion in its accountability system for schools and local educational agencies (LEAs), Educational Test Service evaluated the statistical properties of the California Department of Education’s (CDE’s) three candidate student growth statistics aggregated to the school and LEA levels: *change in distance-to-met* (CDTM), *conditional percentile rank of the gain* (CPR), and *residual gain* (RG). Because the analysis was not intended to evaluate the use of the candidate growth statistics for reporting at the individual student level, our analyses involved within- and across-model comparisons at the aggregate (school or LEA) level and addressed the following three research questions:

*Within-model investigation:*

1. For each model, to what extent do the aggregate growth scores for English language arts/literacy (ELA) and mathematics depend on key demographic variables (e.g., females, English learners [ELs], low socioeconomic status [SES], etc.) at the school level, LEA level, and state level?

*Across-model investigation:*

1. To what extent do aggregate rankings in ELA and mathematics depend on the growth model at the overall school and LEA levels and for student groups within schools and LEAs (e.g., by gender, EL status, SES, etc.)?
2. Which aggregated growth score provides more precise estimates?

Rather than listing the results for each research question one by one, we organize them by key findings to provide a more coherent synthesis of the findings. Similarly, to streamline the presentation of results, we include the most relevant results in the text and additional, supplemental results in the appendixes. We first introduce and define each growth statistic and then review the data used for all analyses. Subsequently, we discuss the relationships of the growth statistics with student demographic variables, followed by the relationships of the growth statistics with school grade configuration. Finally, we evaluate the precision/reliability of the growth statistics. We close with a summary and discussion of the practical implications of the findings.

## Growth Statistics

Each of the three growth statistics the CDE is considering for its accountability system is a function of prior and current test scores. However, they each measure different aspects of student growth and rely on different assumptions. We highlight each in turn.

### Change in Distance-to-Met

California uses (aggregated) distance-to-met (DTM), or the distance from a student’s score to the Met Standard threshold, as the status measure in its accountability system to categorize schools/LEAs by the current performance of their students. Accordingly, change in DTM (CDTM) is the difference in a student’s DTM from one year to the next, as shown in Equation (1) for each content area:



 (1)

where,

the subscripts “1” and “2” denote the prior year as 1 and the current year as 2,

 is the Met Standard threshold for content area *s* in Year *y*, and

 and  denote the ELA and mathematics test scale scores in year *y*, respectively.

To aggregate these statistics to the school/LEA level, we average the CDTMs for all the students enrolled in the school/LEA who have observed CDTM values.

As a concrete example, we consider a student, Johnny, in grade eight. His prior grade seven mathematics score is 2600 and current grade eight mathematics score is 2632. The Met Standard thresholds for grades seven and eight are 2567 and 2586, respectively. Accordingly, Johnny’s DTM in grade seven is 33 score points (2600 - 2567 = 33); he scored 33 points above the grade seven Met Standard threshold. In grade eight, Johnny scored 46 score points above the grade eight Met Standard threshold (grade eight DTM =2632 - 2586 = 46). Thus, he scored 13 more score points above the grade-level Met Standard threshold in grade eight than he did in grade seven (CDTM = 46 – 33 = 13). Johnny is positively increasing the distance between his score and the grade-level Met Standard threshold, indicating better performance relative to the Met Standard threshold in grade eight than grade seven.

Students will have a positive CDTM if they are closer to the Met Standard threshold in the current year than they were in the previous year and scored below the threshold in both years—e.g., a student is 5 points below the Met Standard threshold in the current year versus 10 points below in the previous year: (-5) - (-10) = +5—or if a student exceeds the Met Standard threshold by more in the current year than he or she exceeded or was below the standard in the previous year—e.g., a student is 20 points above the Met Standard threshold in the current year versus 2 points below in the previous year: (20) - (-‍2) = 22. Students whose DTM is the same in both years will have a CDTM of 0; those students who are farther from the Met Standard threshold in the current year than the previous year will have a negative DTM—e.g., a student is 10 points below in the current year versus only 5 points below in the previous year: (-10) - (-5) = -5.

Another way to conceptualize CDTM is that it measures absolute growth on the test score scale against the change in Met Standard thresholds; that is, we can rewrite Equation (1) as shown in Equation (2):

  
 (2)

Equation (2) makes it clear that CDTM is an adjusted *gain score*, where a student’s gain, current year score minus prior year score, is adjusted by the distance between the two years’ Met Standard thresholds. Accordingly, positive CDTMs indicate that a student gained more than the position of the Met Standard changed on the scale, while negative CDTMs indicate that a student gained less relative to the change in Met Standard cuts. A CDTM of zero indicates that a student’s gain is equivalent to the change in Met Standard cuts. The CDTM is a way to anchor the gain score and augment its interpretation.

As an adjusted gain score, CDTM relies on the same assumptions of the test scale as the simple gain score. It relies on a vertical scale with interval scale properties within and across grade levels. Moreover, it relies on placing significant meaning on the location of the Met Standard thresholds in each grade level as we will explore further in the “Relationship with Grade Composition” subsection.

### Conditional Percentile Rank of the Gain Score

The CPR provides a relative or normative measure of student gains. It equals the percentile rank of a student’s gain score relative to peers who had the same prior-year test score in the same content area. Students who gain much more than their peers who started at the same place will have high CPRs (CPRs greater than 50), whereas students whose gains are lower than those who started at the same place will have low CPRs (CPRs less than 50). Students whose gain corresponds to the median gain of their peers with the same prior score will have a CPR of 50.

The percentile rank of the conditional gain is the same as the percentile rank of conditional status when conditional status is defined as current status given a single prior score. That is, because peer comparison groups for CPR are formed by grouping together all students with the same particular prior year score, the percentile rank of a gain score within a given comparison group is equivalent to the percentile rank of the corresponding current status score within the comparison group.

The popular Student Growth Percentile (Betebenner, 2009) statistic is a percentile rank of conditional status but is often computed with several prior scores and uses complicated statistical machinery to approximate the percentile ranks. In contrast, because only one prior year was used and due to the large testing volumes in California, as well as to provide the most straightforward estimation for local replication, we derive the CPRs empirically rather than by introducing statistical modeling assumptions. Specifically, we compute the CPRs for a given grade and content area with the following procedure that is demonstrated using CPRs for grade eight mathematics but is analogous for other grade/content area combinations.

We first define the set of students eligible to be assigned CPRs as those with valid, on-grade mathematics test scores for both grade seven and grade eight. We then partition this set of students into peer comparison groups on the basis of the grade seven mathematics scores, so that there are as many comparison groups as there are unique values of the grade seven mathematics scores. Separately, for each comparison group, we compute the percentile rank of each student’s grade eight mathematics score with respect to his or her comparison group as follows: a student with a grade eight mathematics score of *X* is assigned a percentile rank equal to the percentage of students in the comparison group whose grade eight mathematics score is strictly less than *X* plus one half of the percentage of students in the comparison group whose grade eight score is exactly equal to *X.* That is, each student is given credit for every student whose score he or she exceeds and half a credit for every student he or she ties. This computation of CPRs ensures that the average percentile rank is always 50, both within each comparison group and marginally across all students, regardless of the number of ties in student scores.

Table 1 provides an example for the hypothetical student Johnny introduced in the CDTM section. Johnny scored 2632 on the mathematics test in grade eight and 2600 in grade seven. Among the 1,209 students who scored 2600 on the grade seven mathematics assessment the prior year, i.e., among Johnny’s comparison group, 659 (54.51%) scored below 2632 on the grade eight mathematics assessment in the current year. Another 10 students (0.83%) in Johnny’s comparison group, including Johnny himself, scored exactly 2632. This means that the percentile rank of Johnny’s score among his comparison group is 55, because the percentile rank equals the percentage of students in the comparison group whose grade eight mathematics scores are below Johnny’s score plus half credit for the percentage of the comparison group students who had the same score as Johnny. Johnny’s current grade eight mathematics score is as good as or better than 55 percent of his peers who scored the same as he did in grade seven. Johnny is doing better in grade eight mathematics than more than half of his peers who started at the same place as he did in terms of a grade seven score.

Table 1. Example of computing a mathematics CPR for Johnny, a grade eight student with a prior score of 2600 and a current score of 2632

| Comparison Group: Number of students with prior score = 2600 | Percentage of students in “2600” comparison group with current score < 2632 | Percentage of students in “2600” comparison group with current score = 2632 | CPR |
| --- | --- | --- | --- |
| 1,209 | 54.51% | 0.83% |  |

The CPR relies on the fewest assumptions of the test scale of the three considered growth statistics. It does not require a vertical scale or that the within-grade or across-grade scales have interval properties. Some students, however, may have imprecisely estimated CPRs if there are few or no other students with the same prior score. In the cases where a student has no peers with the same prior score, the student will receive a CPR of 50 (the average). Because the California ELA and mathematics assessments are vertically scaled, there are many distinct score points in any given grade and thus, some instances where this issue occurs.

Another consideration with CPRs, unlike with CDTMs (which can be estimated for a particular student with the student’s scores alone), is that CPRs require having data for the full state available before they can be estimated.

### Residual Gain

The RG, like CPR, is a measure of relative or normative growth. It is used to answer the question: How much better/worse did a student do this year than expected, given the student’s prior scores? The expected performance is estimated using a linear regression. For this study, the interest is in conditioning on both ELA and mathematics prior scores, as shown in Equation (3):

 (3)

where,

 denotes the expected current ELA score given the prior ELA and mathematics scores, and likewise,  denotes the expected current mathematics score given the prior ELA and mathematics scores;

 and  are the intercepts in the ELA and mathematics regressions, respectively; and

 are the regression coefficients for the prior scores in each regression.

The residual gains are the difference between a student’s actual current score and the student’s expected current score, as shown in equation (4):

  (4)

Accordingly, positive RGs indicate a student scored higher than expected given the student’s prior performance and vice versa for negative RGs.

As a concrete example, we consider Johnny’s performance again. On the basis of his prior year grade seven mathematics and ELA scores of 2600 and 2610, Johnny’s expected grade eight mathematics score is 2629.12. Because Johnny actually scored 2632 on the grade eight mathematics assessment, his grade eight mathematics RG is 2.88 (‍2632 - 2629.12). Johnny scored about 3 score points higher in grade eight mathematics than the average student with the same prior grade seven ELA and mathematics scores as Johnny.

RGs do not require a vertical scale, but they do require an interval scale within grade. Like CPRs, they require having the full state of data available to estimate them.

## Data

For all analyses, we used a longitudinal, student-level dataset of CAASPP ELA and mathematics test scores from the 2014–15 and 2015–16 school years. ETS used a file provided by the CDE Accountability office that only included students with valid test scores (i.e., continuously enrolled students). The data included student identifiers as well as school and LEA identifiers, several student demographic variables for the current year (see Table A.1 in Appendix A for a full list of student variables used in the analyses), and the charter-funding status of schools (school-level variable).

The data required minor cleaning before we could estimate student growth scores and aggregate those scores to the school and LEA levels. Following discussions with the CDE, we removed cases with the condition code of lowest obtainable scale score (LOSS) in the current or prior year[[3]](#footnote-4), which resulted in the very minimal reduction of 37 to 141 students per grade level for each content area out of roughly 410,000 students per grade. We also made the following manual changes to student covariates to ensure consistency across student records in mathematics and ELA data files:

1. There was one student missing gender, with only ELA scores observed. Because males are slightly more prevalent in the sample, we classified this student as male.
2. There were 40 students whose race was discrepant between the mathematics and ELA files. We classified all of these students with code 800 corresponding to “Two or more races.”
3. There were 19 students whose SWD (student with disability) designation was discrepant between the mathematics and ELA files, meaning that they were classified as SWD in exactly one of the two files. We coded such students as having a disability (SWD=1).
4. There were 12 students whose SED (socioeconomically disadvantaged) designation is discrepant between the mathematics and ELA files, meaning that they were classified as SED in exactly one of the two files. We coded such students as being socioeconomically disadvantaged (SED=1).

To estimate the CDTMs and CPRs for students in each content area, students needed to have prior and current scores in the content area of interest. To estimate the RGs, students needed to have the current score in the content area of interest in addition to having prior scores in both content areas making the data requirements for RGs slightly more stringent than for the other two measures. However, as seen in Table 2 and Table 3, less than one percent of students were lost for RG for school- and LEA-level analyses. All students were used to estimate student growth statistics, but only schools and LEAs with a minimum of 10 students were used in subsequent analyses. Table 2 and Table 3 also summarize the loss in number of students and schools/LEAs by imposing this constraint; close to zero percent of students were lost and about two percent of schools and four percent of LEAs were lost. There are fewer students included in the LEA-level analysis than the school-level analysis because students in charter schools are not aggregated to the authorizing LEA for purposes of accountability.

Table 2. Summary of data used for analyses—School level

| Data Category | ELA CPR and CDTM | ELA RG | ELA % Loss for RG | Math CPR and CDTM | Math RG | Math % Loss for RG |
| --- | --- | --- | --- | --- | --- | --- |
| All students | 2,059,971 | 2,047,548 | 0.60% | 2,055,236 | 2,04,7521 | 0.38% |
| Students in schools ≥ 10 | 2,059,390 | 2,046,968 | 0.60% | 2,054,653 | 2,046,942 | 0.38% |
| %Loss | 0.03% | 0.03% | No data | 0.03% | 0.03% | No data |
| All Schools | 7,466 | 7,466 | 0.00% | 7,466 | 7,466 | 0.00% |
| Schools ≥ 10 students | 7,327 | 7,324 | 0.04% | 7,325 | 7,324 | 0.01% |
| %Loss | 1.86% | 1.90% | No data | 1.89% | 1.90% | No data |

Table 3. Summary of data used for analyses—LEA level

| Data Category | ELA CPR and CDTM | ELA RG | ELA % Loss for RG | Math CPR and CDTM | Math RG | Math % Loss for RG |
| --- | --- | --- | --- | --- | --- | --- |
| All students | 1,883,540 | 1,872,796 | 0.57% | 1,879,443 | 1,872,776 | 0.35% |
| Students in  LEAs ≥ 10 | 1,883,393 | 1,872,653 | 0.57% | 1,879,299 | 1,872,633 | 0.35% |
| %Loss | 0.01% | 0.01% | No data | 0.01% | 0.01% | No data |
| All LEAs | 860 | 860 | 0.00% | 860 | 860 | 0.00% |
| LEAs ≥ 10 students | 828 | 828 | 0.00% | 828 | 828 | 0.00% |
| %Loss | 3.72% | 3.72% | No data | 3.72% | 3.72% | No data |

Notes:

* “% Loss” is the percentage of students, schools, or LEAs excluded from the analysis by the requirement that schools or LEAs included in the analysis have 10 or more students in the sample.
* “% Loss for RG” is the percentage of students, schools, or LEAs excluded from the analysis by the requirement for the residual gain (RG) that students have prior scores in both content areas.

Some analyses involved the mean growth statistic for student groups within a school or LEA (e.g., mean growth for females within a school). In those cases, we restricted the dataset to those schools/LEAs with any records for at least 10 students in the student group of interest.

Having described the three growth statistics and the data, we now move to describing analyses we performed to compare the growth statistics. The subsections in the “Analyses” section discuss:

1. the relationships of the growth statistics with student demographic variables,
2. the relationships of the growth statistics with school grade configuration, and
3. the precision/reliability of the growth statistics.

## Analyses

### Relationships with Student Demographic Variables

To assess the relationships between each growth statistic for schools and LEAs and aggregate student demographic characteristics, we obtained two sets of statistics: (1) correlations between each growth statistic and prior mean scores in each content area as well as with each aggregated student demographic variable; and (2) the values from first regressing the school/LEA mean growth statistic on both the mean prior scores and aggregated student demographic characteristics, and then from regressing on only the aggregated student demographic characteristics to determine the extent that the variance in the school/LEA mean growth statistics is explained by the demographic composition of the students in the schools/LEAs. Before estimating each set of outcome statistics, we transformed the aggregated student demographic characteristics (e.g., proportion of EL students) with the arcsine square root normalizing transformation because the distributions of the proportions are often skewed, with few schools having large proportions of students in a particular student group.

Table 4 (with Figure 1 and Figure 2) and Table 5 (with Figure 3 and Figure 4), for school and LEA levels, respectively, present the results of this analysis. The figures provide the correlations with each (transformed) aggregated student demographic variable (defined in Table A.1) with the variables, ordered by the magnitude of the correlation between RG and each variable to enhance readability of the figures. Table B.1 through Table B.4 provide the data values plotted in Figure 1 through Figure 4.

The table and figures clearly show that the selection of the growth measure will affect the relationship between school or LEA growth measures and the characteristics of the students that they serve. For example, at the school level, CDTM has a weak negative relationship with average prior achievement for ELA, but the other growth measures are positively related with average prior achievement. Also, CPR tends to have stronger relationships with student covariates as shown by the high values in the last column of Table 4 and Table 5 *(**for Mean Student Covariates Only)* and in the figures by the values for CPR often being further from zero than those of the other two growth measures.

Table 4. Relationships of each growth statistic with average prior scores and student covariates—school level

| Content Area | Growth Statistic | Correlation with Prior Mean ELA | Correlation with Prior Mean Math | for Mean Student Covariates + Mean Prior Scores | for Mean Student Covariates Only |
| --- | --- | --- | --- | --- | --- |
| ELA | CDTM | -0.13 | -0.07 | 0.10 | 0.03 |
| ELA | CPR | 0.24 | 0.30 | 0.18 | 0.15 |
| ELA | RG | 0.13 | 0.11 | 0.08 | 0.08 |
| Mathematics | CDTM | 0.40 | 0.38 | 0.17 | 0.12 |
| Mathematics | CPR | 0.26 | 0.28 | 0.17 | 0.17 |
| Mathematics | RG | 0.17 | 0.23 | 0.14 | 0.12 |

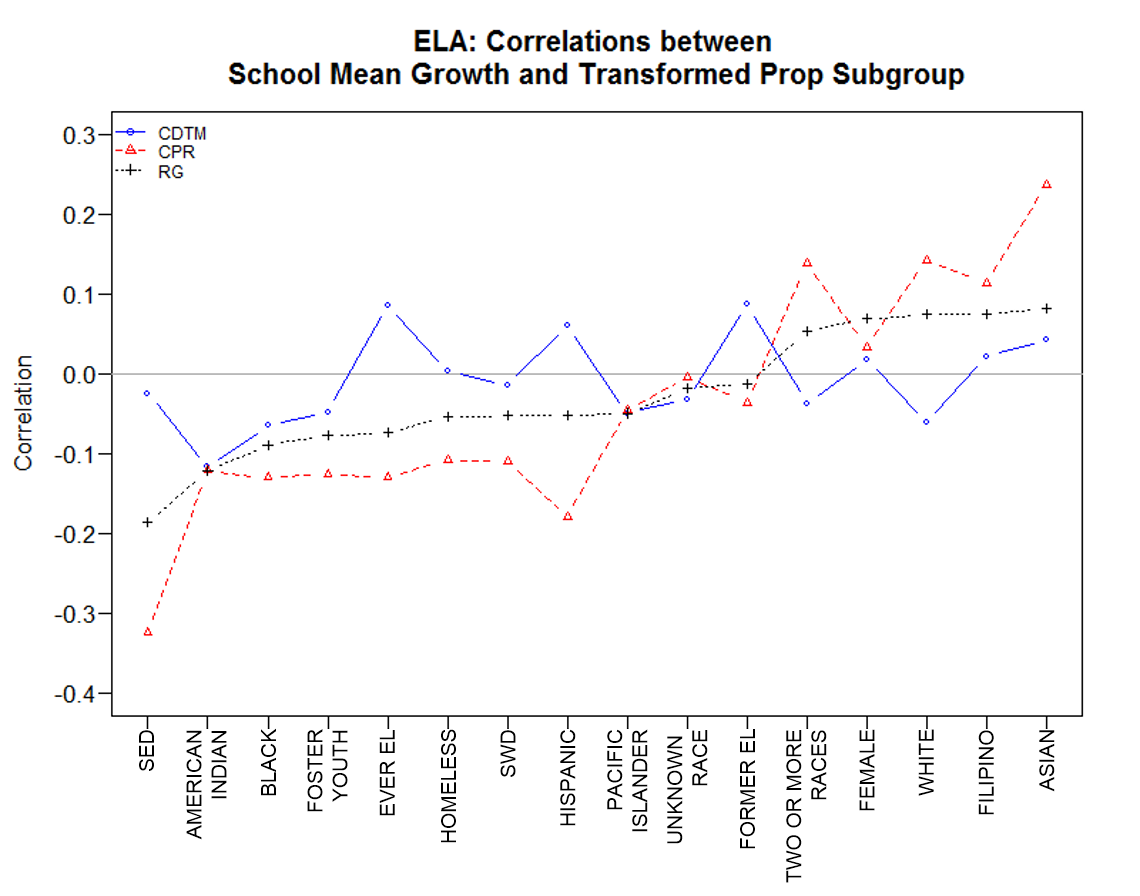


Figure 1. Correlations between each mean growth statistic at the school level for ELA and each transformed aggregate student demographic variable.

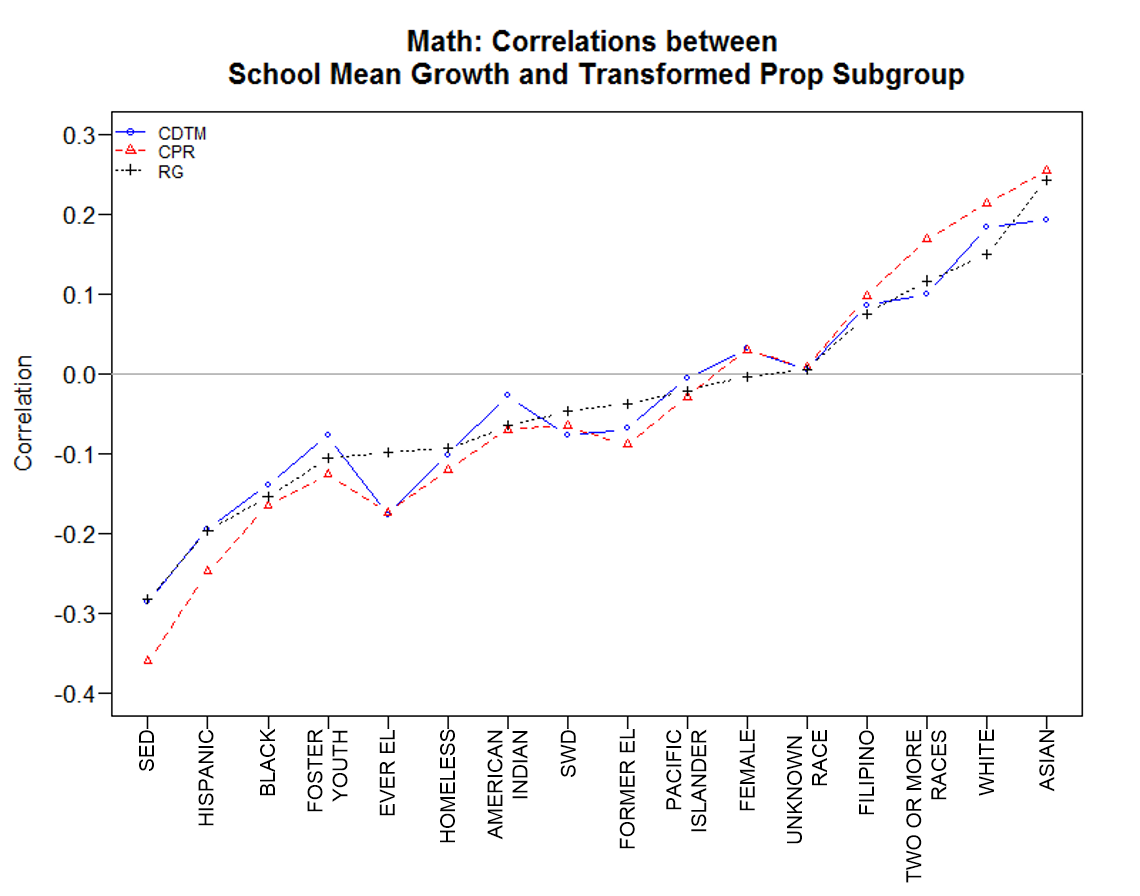


Figure 2. Correlations between each mean growth statistic at the school level for mathematics and each transformed aggregate student demographic variable.

Table 5. Relationships of each growth statistic with average prior scores and student covariates—LEA level

| Content Area | Growth Statistic | Correlation with Prior Mean ELA | Correlation with Prior Mean Math | for Mean Student Covariates + Mean Prior Scores | for Mean Student Covariates Only |
| --- | --- | --- | --- | --- | --- |
| ELA | CDTM | 0.01 | 0.04 | 0.11 | 0.08 |
| ELA | CPR | 0.39 | 0.41 | 0.25 | 0.24 |
| ELA | RG | 0.24 | 0.19 | 0.17 | 0.16 |
| Mathematics | CDTM | 0.37 | 0.35 | 0.21 | 0.20 |
| Mathematics | CPR | 0.40 | 0.39 | 0.26 | 0.25 |
| Mathematics | RG | 0.29 | 0.33 | 0.21 | 0.19 |

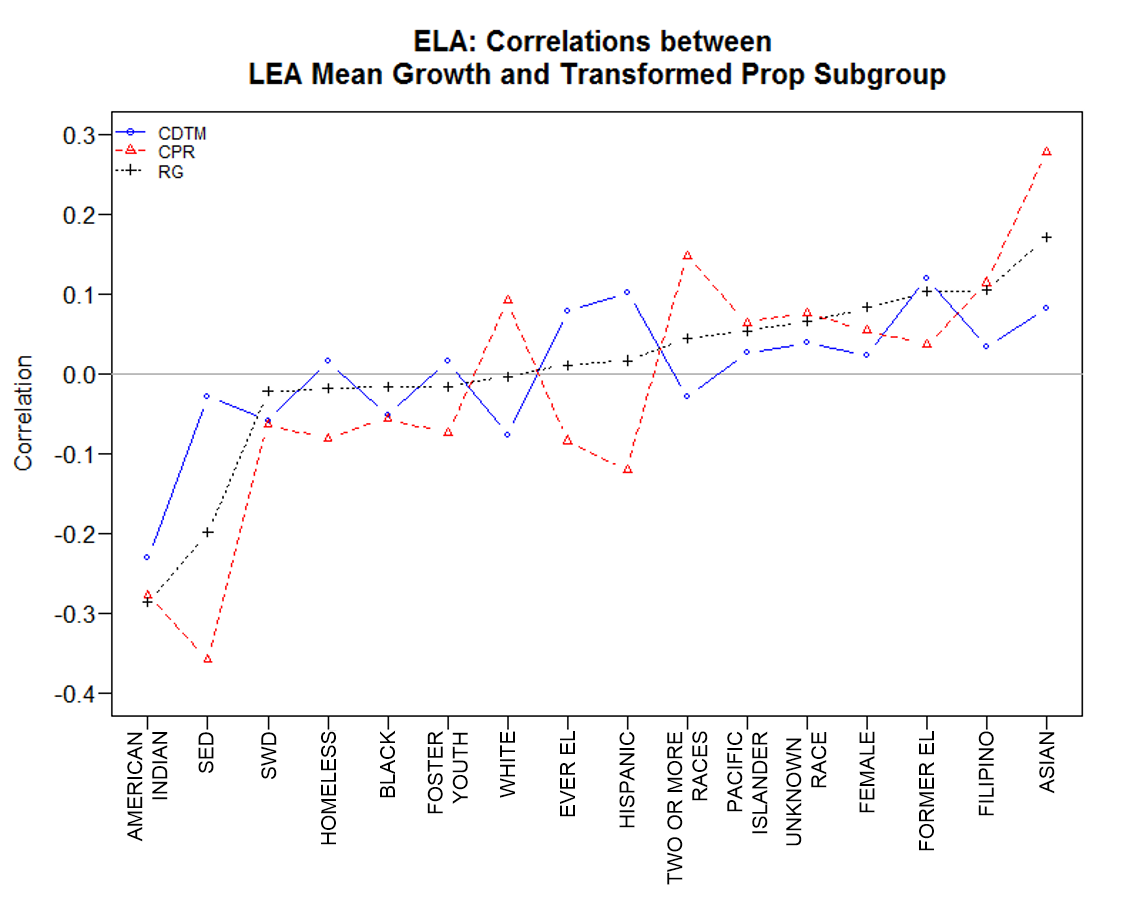


Figure 3. Correlations between each mean growth statistic at the LEA level for ELA and each transformed aggregate student demographic variable.

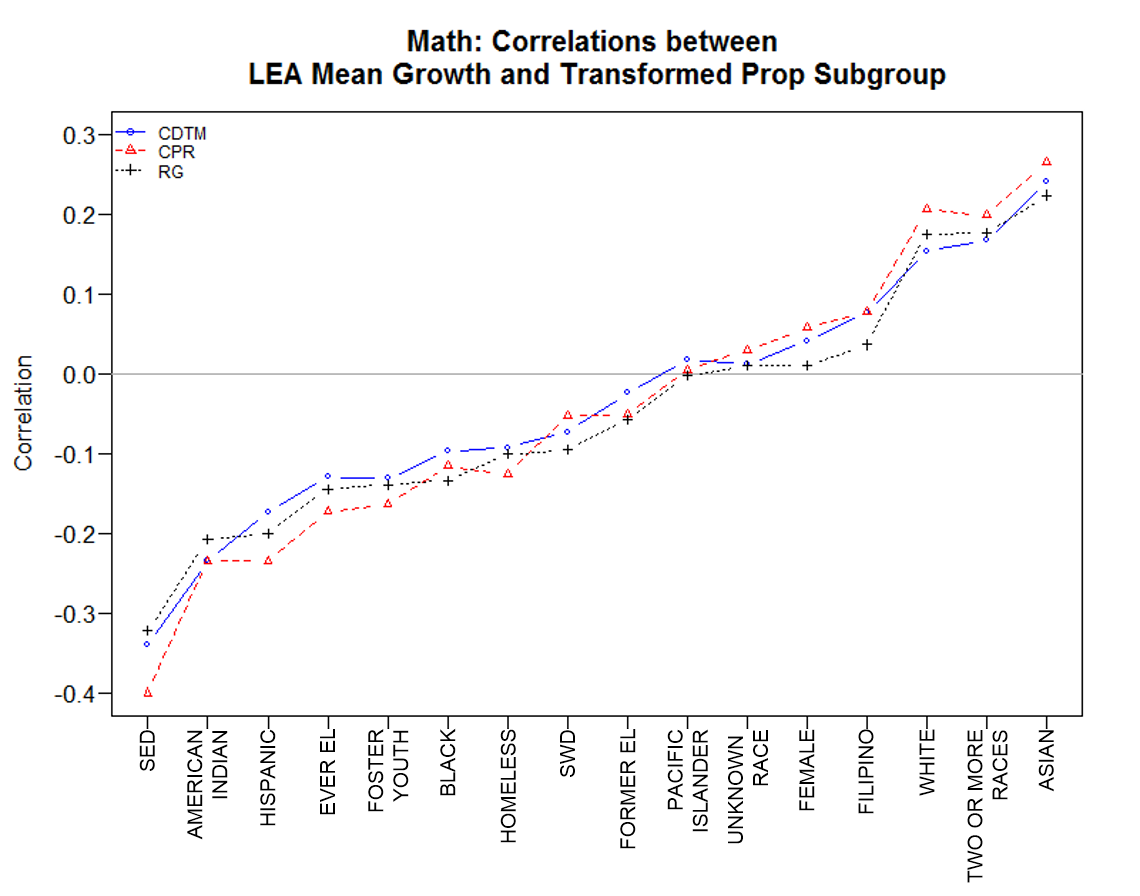


Figure 4. Correlations between each mean growth statistic at the LEA level for mathematics and each transformed aggregate student demographic variable.

Although these relationships are generally weak to modest, they can result in notable differences for schools serving distinctly different populations of students. To illustrate these notable differences, we consider ranking the schools according to each of the three candidate growth measures and examine how different the average percentile rank is between schools serving either a large or small percentage of socioeconomically disadvantaged (SED) students. Specifically, for each growth measure, we compare the average school percentile rank for schools primarily serving socioeconomically disadvantaged students (≥ 75% of SED students) to the average for schools with a low incidence of such students (≤ 25% of SED students). There are 1,144 low-SED schools and 3,482 high-SED schools with at least 10 students for each growth measure.

As seen in Figure 1, the correlations between the school aggregate growth measure for ELA and (transformed) percentage of SED students are near 0 (≈ -.03) for CDTM, about -‍0.19 for RG, and about -0.32 for CPR. Accordingly, as shown in Table 6, we find little difference between the mean percentile ranking of schools by ELA CDTM serving these two distinct populations, but the modest negative correlation of -0.32 for ELA CPR results in a difference of 27 percentile points, on average, between low and high SED schools. Specifically, schools where 25 percent or less of their students are classified as SED are ranked at the 69th percentile on average by their school mean ELA CDTM value while those with 75 percent or more SED students are ranked at the 42nd percentile on average. The correlations for the school mean growth measures in mathematics are more similar (see Figure 2), ranging from -0.28 to -0.36, yielding similar differences in percentile ranking by all three measures for the two distinct sets of schools. All of these mean rank differences are very large, with that for CPR again being the largest.

Table . Average percentile rank by school mean growth for each of the growth measures of interest for schools with low percentage (≤ 25%) of socioeconomically disadvantaged students versus a high percentage (≥ 75%)

| Percentage SED | ELA CDTM | ELA CPR | ELA RG | Math CDTM | Math CPR | Math RG |
| --- | --- | --- | --- | --- | --- | --- |
| ≤ 25% | 53 | 69 | 60 | 65 | 70 | 67 |
| ≥ 75% | 50 | 42 | 45 | 42 | 41 | 43 |
| Difference | 3 | 27 | 14 | 22 | 29 | 24 |

Appendix A provides figures comparable to Figures 1–4 for correlations between school/LEA student group mean growth statistics and each (transformed) aggregate student demographic variable as well as boxplots depicting the full distributions of the student group school/LEA means for each growth statistic. Each figure is followed with a data table of the plotted values.

Differences among the growth measures are consistent with those in Table 4 and Table 5 and their related figures. For instance, the correlations with aggregate student covariates are again generally larger for CPR than the other measures.

#### Interpreting Differences in the Relationships with Student Demographic Variables

The goal of calculating school/LEA level aggregated student growth measures, typically, is to obtain information about how well a school is promoting student achievement growth on the constructs measured by the statewide assessment. Growth is often used in addition to current achievement status because growth is a measure of learning and is generally believed to be less strongly influenced by student demographic variables than status measures (e.g., current scale score). Status measures are valuable for determining whether individual schools’ students are “getting where they need to be,” but they may disadvantage schools serving traditionally disadvantaged populations. Growth is meant to better level the playing field and provide information about what a school contributes annually to a student’s learning. Accordingly, if all else is equal, when comparing two growth measures, one that is less conflated with student demographic information would be preferred for helping inform schools about their performance.

Among the three growth statistics of interest, CPR appears to fall behind the other two on such a comparison. The problem with CPR is due to measurement error in the prior-test score and the negative impact this has on the ability of conditioning on the prior-test score to create truly comparable sets of students (see, for instance, McCaffrey, Castellano, & Lockwood, 2015, for a discussion of measurement error bias in the accuracy of student- and group-level student growth percentiles). To describe the problem, we first consider how the measures behave at the student level and across subpopulations of students with different demographic characteristics. We then evaluate the implications for the behavior of the aggregate measures.

Suppose there is a dichotomous grouping variable G representing some student demographic variable such as race/ethnicity (e.g., White versus non-White) or SED (0 versus 1). Denote the two disjoint populations of students implied by the variable G by G1 and G2. Suppose that students in G1 have, on average, higher true prior achievement than students in G2. Then, under reasonable assumptions about the nature of the measurement error in the prior-test score, it is the case that students in G1 with a particular prior, observed test score *X* still will have, on average, higher true prior achievement than the corresponding students in G2 with the same prior, observed test score *X*. This results from the fact that aligning students on the basis of the error-prone, observed test score *X* is not the same as aligning students on the basis of true (unobserved) prior achievement. The higher the reliability of the prior score, the less this problem manifests.

If students in G1 and G2 have, on average, truly equal growth, then a growth statistic that is computed by first conditioning on the prior-test score and then examining growth within comparison groups formed on the basis of the prior-test score will suggest that students in G1 had higher growth on average than students in G2. This is because the growth statistic contains an echo of the pre-existing achievement differences between the groups. More generally, any differences in true growth between the groups will be, to some extent, confounded with prior achievement status differences between the groups. Again, this issue becomes less impactful as the reliability of the prior score increases.

From a purely analytical standpoint, it is expected that this problem should manifest more severely for CPR than it does for CDTM or RG. Because CDTM relies on gain scores rather than conditioning, it is capable, under appropriate assumptions, of completely removing prior average achievement differences between different groups of students. Alternatively, RG is still susceptible to the same problem as CPR, but the problem generally should be less pronounced with RG than it is for CPR because RG conditions on prior scores from both mathematics and ELA. Because true prior mathematics and ELA achievement are relatively strongly correlated, conditioning on both prior-test scores can act like a type of “correction” for the problem caused by measurement error in each individual prior score. It is worth noting that the use of the percentile rank scale for CPR does not change these basic arguments.

To help illustrate that these theoretical concerns appear to manifest in the CAASPP data, we created two additional growth measures from the available data: “RG1,” or the residual gain computed by regressing on only the same-content area prior score; and “RG1adj,” or the residual gain computed by regressing on the same-content area prior score with an adjustment for its measurement error in the prior score. The adjustment for measurement error in the prior score was conducted with standard procedures (i.e., errors-in-variable regression, Fuller, 2006; also see Culpepper & Aguinis, 2011, and references therein) using information about the conditional standard errors of the test scores provided in the data.

The theoretical considerations suggest that CPR and RG1 should behave very similarly because they both condition on only the same-content area prior score, and they should, on average, imply larger differences in growth between groups of students with different demographic characteristics than either RG or RG1adj. See for example, Lockwood and McCaffrey (2014) for a discussion of how including multiple prior-test scores or correcting for measurement error affects estimates of classroom average growth estimated via regression models. RG and RG1adj approach the measurement error problem differently: RG only implicitly corrects for measurement error by using two prior-test scores in the conditioning, whereas RG1adj directly corrects for measurement error in the same-content area prior score using auxiliary information about the magnitude of the measurement error. However, both statistics should tend to imply smaller differences in growth, compared to either CPR or RG1, between groups of students with different demographic characteristics.

Empirical analyses are largely consistent with these *a priori* expectations. Table 7 presents contrasts in average ELA growth measures between various student groups formed on the basis of student demographic variables. Different columns correspond to different growth measures, and different rows correspond to different contrasts. Prior to computing the contrasts, each growth measure was standardized to have mean zero and variance one in the population. Thus, the contrasts are in standard deviation units of each growth measure. For example, the *White–Black* row indicates that White students had an average CDTM growth statistic that is 0.07 population standard deviation units of the CDTM growth statistic larger than the average for Black students, and White students had an average CPR growth statistic that is 0.21 population standard deviation units of the CPR growth statistic larger than the average for Black students.

Table 7. Average standardized group differences for each growth statistic in ELA

| Demographic Student Group Comparison | CDTM | *RG1adj* | RG | *RG1* | CPR |
| --- | --- | --- | --- | --- | --- |
| Female–Male | 0.06 | 0.08 | 0.23\* | 0.12 | 0.12 |
| White–American Indian | 0.05 | 0.08 | 0.13 | 0.19 | 0.17\* |
| White–Asian | -0.10 | -0.12 | -0.06 | -0.16 | -0.18\* |
| White–Pacific Islander | 0.01 | 0.04 | 0.08 | 0.12 | 0.11\* |
| White–Filipino | -0.07 | -0.07 | -0.09\* | -0.09 | -0.08 |
| White–Hispanic | -0.01 | 0.03 | 0.07 | 0.13 | 0.12\* |
| White–Black | 0.07 | 0.11 | 0.13 | 0.23 | 0.21\* |
| White–Two or more races | -0.01 | 0.00 | -0.01 | -0.01 | -0.02\* |
| White–Unknown (Missing) | 0.01 | 0.01 | 0.02 | 0.04 | 0.04\* |
| Never EL–Current EL | 0.01 | 0.07 | 0.20 | 0.24 | 0.22\* |
| Never EL–Past EL | -0.06\* | -0.06 | -0.04 | -0.03 | -0.03 |
| Never EL–Ever EL | -0.03 | 0.01 | 0.08 | 0.10 | 0.09\* |
| Not SWD–SWD | 0.13 | 0.18 | 0.26 | 0.34 | 0.33\* |
| Not SED–SED | 0.03 | 0.07 | 0.12 | 0.20 | 0.19\* |
| Not homeless–Homeless | 0.01 | 0.04 | 0.07 | 0.11 | 0.10\* |
| Not foster youth–Foster youth | 0.06 | 0.10 | 0.14 | 0.20 | 0.20\* |
| Mean Absolute Difference | 0.04 | 0.07 | 0.11 | 0.14 | 0.14 |

Notes:

* The largest difference in absolute value among the three primary growth statistics (CDTM, CPR, and RG) is marked with an asterisk (\*) in each row.
* RG1adj and RG1 are included to help demonstrate the impact of measurement error on group comparisons. They are not among the growth measures being evaluated in this report.

Aside from a few exceptions (e.g., the Female–Male contrast, which may behave differently because average status scores between Females and Males are less different than for other contrasts considered), CPR and RG1 imply similar group differences in growth that tend to be larger than the other measures. As shown by the values marked with an asterisk (\*) in Table 7, CPR has larger group differences than either of the other two primary growth statistics of interest (RG and CDTM) in the majority of group comparisons. As expected, RG generally behaves between RG1 and RG1adj (i.e., the values for RG in Table 7 are typically between those for RG1 and RG1adj), which is likely to result from the fact that RG is known to be a partial, rather than a complete, correction for measurement error. CDTM has on average the smallest group contrasts, primarily because it does not use direct conditioning on the error-prone prior-test scores.

The patterns for the mathematics growth statistics in Table 8 are very similar to those for ELA, although RG1adj has the smallest absolute differences on average for mathematics, which is in contrast to CDTM for ELA. Appendix C provides analogous results, with figures, for comparing student group means at the school and LEA levels.

Table 8. Average standardized group differences for each growth statistic in mathematics

| Demographic Student Group Comparison | CDTM | *RG1adj* | RG | *RG1* | CPR |
| --- | --- | --- | --- | --- | --- |
| Female–Male | 0.03 | 0.03 | -0.08\* | 0.03 | 0.03 |
| White–American Indian | 0.15 | 0.10 | 0.14 | 0.20 | 0.17\* |
| White–Asian | -0.16 | -0.11 | -0.22\* | -0.18 | -0.19 |
| White–Pacific Islander | 0.09 | 0.05 | 0.08 | 0.12 | 0.10\* |
| White–Filipino | -0.08\* | -0.06 | -0.06 | -0.07 | -0.07 |
| White–Hispanic | 0.13 | 0.06 | 0.12 | 0.16 | 0.14\* |
| White–Black | 0.20 | 0.13 | 0.22 | 0.25 | 0.24\* |
| White–Two or more races | 0.02\* | 0.00 | 0.00 | 0.00 | 0.00 |
| White–Missing | 0.03 | 0.01 | 0.03 | 0.04 | 0.04\* |
| Never EL–Current EL | 0.19 | 0.07 | 0.09 | 0.22 | 0.21\* |
| Never EL–Past EL | -0.01 | -0.04 | -0.03 | -0.01 | -0.03\* |
| Never EL–Ever EL | 0.08\* | 0.02 | 0.03 | 0.10 | 0.08\* |
| Not SWD–SWD | 0.18 | 0.09 | 0.15 | 0.23 | 0.27\* |
| Not SED–SED | 0.17 | 0.10 | 0.16 | 0.22 | 0.20\* |
| Not homeless–Homeless | 0.09 | 0.04 | 0.07 | 0.10 | 0.10\* |
| Not foster youth–Foster youth | 0.14 | 0.08 | 0.14 | 0.18 | 0.18\* |
| Mean Absolute Difference | 0.11 | 0.06 | 0.10 | 0.13 | 0.13 |

Notes:

* The largest difference in absolute value among the three primary growth statistics (CDTM, CPR, and RG) is marked with an asterisk (\*) in each row.
* RG1adj and RG1 are included to help demonstrate the impact of measurement error on group comparisons. They are not among the growth measures being evaluated in this report.

The correspondence of the empirical results with the theoretical considerations suggests that CPR tends to have stronger spurious relationships with student demographics than the alternative measures. The use of CPR will likely result in inferences about the relative growth of different student groups that are exaggerated relative to the corresponding true differences in growth.

The spurious differences among student groups for student-level growth caused by measurement error in the prior scores also manifest for the school- and LEA-level aggregate growth statistics. Again, CPR fares relatively worse than either RG or CDTM in terms of introducing an overall larger magnitude of spurious differences in the growth measures among schools or LEAs relative to the alternatives. This is because measurement error in the prior-test scores makes the adjustment for a single prior-test score insufficient to control for prior achievement differences among students in different schools (or LEAs). If there are two schools (“A” and “B”) that are equally effective at raising student achievement in the current year, but school A serves a student population with lower average true prior achievement than school B, school A’s average growth on a measure such as CPR will tend to be lower than school B’s. More generally, there will be a source of spurious positive correlation between schools’ growth measures and schools’ average prior achievement. Such a problem will be worse for a growth measure that does a poorer job of adjusting for true prior achievement differences. Similarly, there will be a source of spurious correlation between schools’ growth measures and school aggregate demographic (background) characteristics for any such characteristics that are correlated with true prior achievement. That is the case for most, if not all, of the demographic variables under consideration (e.g., race/ethnicity, SED, EL, SWD, etc.) and thus is of interest to this study.

These theoretical considerations are consistent with the patterns in Table 4 and Table 5 and their respective figures. Table 4 and Table 5 show that the correlations between each school/LEA growth score and prior mean scores are always higher for CPR than RG; likewise, so are the  values. For instance, school-mean ELA CPR is correlated .24 with prior mean ELA scores, but school-mean ELA RG is only correlated .13 with prior mean ELA scores. The corresponding correlations for RG1 and RG1adj are .25 and .04, respectively, with CPR closer to that for RG1 and RG closer to that for RG1adj. Similarly, for school-level mathematics results, the correlations with prior mean mathematics scores are .28 for CPR, .29 for RG1, .23 for RG, and .11 for RG1adj, which again show that CPR behaves more like RG1 and RG more like RG1adj. We find comparable results at the LEA-level and for the  values. Also as noted previously, Figures 1–4 show that correlations with aggregate student covariates are generally larger in absolute value for CPR (red dashed lines) than for RG (black dotted lines) for each content area and level (school or LEA).

Measurement error in the prior score can also create a spurious relationship between prior scores and gain scores and, consequently, CDTMs, which are a function of gain scores. This effect might be part of the reason for the negative correlation between CDTM and prior achievement for ELA at the school level. If there is no relationship between prior achievement and gains in achievement in ELA, then the spurious negative relationship due to measurement error could result in a negative correlation between CDTM and prior-test scores.

### Relationship with Grade Composition

In this analysis we study how the three targeted growth measures (CDTM, CPR, and RG) differ as a function of the grade configuration of schools. For accountability at the school or LEA level, student growth statistics are aggregated across grade levels within each content area.

Schools have students in very different compositions of grades. At the LEA level, there is not as much variability in grades represented because each LEA is often a combination of several schools serving different grade levels, so over all the schools together, most of the grade levels of interest—grades four to eight—are represented. In fact, in our data, about 87 percent of the LEAs in California serve students across the full range of grades four to eight. Accordingly, our focus in this section is at the school level, which has a total of 28 different grade configurations represented among grades four to eight in the 2015–16 academic year. Table 9 lists the top 12 represented grade configurations.

Table 9. Top 12 grade configurations for schools

| Grade Configuration | Average Grade | Frequency | Percent\* |
| --- | --- | --- | --- |
| 4, 5 | 4.5 | 2572 | 34% |
| 4, 5, 6 | 5 | 2060 | 28% |
| 4, 5, 6, 7, 8 | 6 | 1063 | 14% |
| 6, 7, 8 | 7 | 962 | 13% |
| 7, 8 | 7.5 | 452 | 6% |
| 4 | 4 | 86 | 1% |
| 5, 6, 7, 8 | 6.5 | 85 | 1% |
| 4, 5, 6, 7 | 5.5 | 47 | 1% |
| 8 | 8 | 20 | 0% |
| 4, 6, 7, 8 | 6.25 | 13 | 0% |
| 6, 7 | 6.5 | 13 | 0% |
| 4, 5, 6, 8 | 5.75 | 10 | 0% |

\*Note: The percent is found by dividing the frequency by the total number of schools (N=7,466).

To investigate the extent each aggregated school growth statistic depends on the school’s grade configuration, we correlated each measure with the average grade level that the school serves. Average grade equals the simple average of the available grade levels of students at each school. For instance, as shown in Table 9, if a school has growth scores for students in grades four, five, and six, then the school’s average grade level is five.

Table 10 provides these correlations in the top two rows. It shows that both CPR and RG have weak, negative relationships with average grade in both ELA and mathematics (*r* ≈ -0.11). However, CDTM has a stronger, negative relationship with average grade level for ELA with a correlation of -0.25 and a stronger, positive relationship for mathematics with a correlation of +0.23. Thus, CDTM has stronger relationships with school grade configuration than the other two measures and the direction of the relationship differs markedly by content area.

Table 10. Correlations to establish relationships with each school growth measure and grade configurations by content area

| Correlation | CDTM | CPR | RG |
| --- | --- | --- | --- |
| ELA growth & Average Grade | -0.25 | -0.11 | -0.11 |
| Mathematics growth & Average Grade | 0.23 | -0.11 | -0.13 |
| (Mathematics–ELA growth) & Average Grade | 0.48 | -0.02 | -0.03 |
| ELA growth & Mathematics growth | 0.49 | 0.69 | 0.63 |

#### Interpreting Differences in Relationships with Grade Composition

For an accountability system, it is not typically desirable to have a school’s growth statistic aggregated across grades depend on which grade levels of students the school happens to serve. One of the original motivations for considering measures other than simple gain scores is that average gain scores vary by grade level, with average gains tending to decrease as the grade increases. Accordingly, school mean gain scores will tend to have strong sensitivity to schools’ grade configurations. As seen in this study, school average gain scores are correlated nearly -0.6 for ELA and nearly -0.5 for mathematics with average grade level. The strong variation of CDTM across grade levels, therefore, could be undesirable.

CDTM’s behavior follows from the fact that it does not directly remove variation across grades in average gains. Rather, it attempts to remove variation indirectly by adjusting for differences in Met Standard cut points across grades. However, the threshold locations in their respective distributions vary across both grades and content areas, which makes the patterns across grades look extremely different for mathematics and ELA, as seen in the top panel of Figure 5. This panel plots the percentile rank of each Met Standard threshold by grade level and content area. For instance, in grade four, the Met Standard threshold is at about the 55th percentile for ELA (solid line, open circles), while it is just above the 60th percentile for mathematics (dashed line, closed circles). The location of the cuts are almost mirror images of each other by content area: when the threshold moves up in the distribution for one content area, it moves down for the other content area. The bottom panel shows the differences in threshold location between adjacent grade levels for each content area. The lines cross, illustrating the contrasting patterns in threshold location by content area across grade levels.

Table B.5 and Table B.6 in Appendix B contain the data values plotted in Figure 5.

Percentile of Met Standard thresholds for each grade level and content area. See previous paragraph for a description of Figure 3. This graph uses data from Table B.5

**Difference in percentiles of thresholds for each pair of adjacent grade levels. See previous paragraph for a description of Figure 3. This graph uses data from Table B.6.**

Figure 5. Percentiles of Met Standard thresholds for each grade level and content area (top panel) and differences in percentiles of thresholds for each pair of adjacent grade levels (bottom panel).

Given the interaction of threshold locations by content area, there is a strong interaction by content area with average school grade level for CDTM, where there is a positive relationship for mathematics and a negative relationship for ELA. Thus, not only are the magnitudes of the relationships larger than for CPR and RG, there is a strong interaction by content area that does not exist for the other measures.

An implication of this content area interaction for CDTM is that the difference between the mathematics and ELA measures for a school are rather strongly related to school grade configuration (correlation near 0.5), whereas the corresponding correlations for the other measures are both near zero, as shown in the *(Mathematics–ELA growth) & Average Grade* row of Table 10. Even simple average gains have a correlation of about only 0.15 (not shown in table). More generally, the mathematics and ELA CDTM are more weakly correlated at the school level than any of the other measures, as shown in the *ELA growth & Mathematics growth* row of Table 10. This means there will be more schools whose average growth measures are distinctly different for the two content areas under CDTM than any of the other measures. Accordingly, schools that serve students in higher grade levels, on average, are more likely to have a larger difference in their mathematics and ELA CDTM values. CDTM demonstrates grade and content area sensitivities that require one to believe that the DTM criterion and location of thresholds are truly meaningful.

### Precision/Reliability

Aggregated growth statistics are susceptible to sampling error. Thus, it is useful to provide a measure of their precision and compare the precision across the candidate growth measures.

For each growth measure and content area, we estimated standard errors for both the school-level and LEA-level aggregates. We used similar procedures to calculate these standard errors for mathematics and ELA, for the different growth measures, and for both levels of aggregation. In broad terms, the standard error for the aggregate growth measures was obtained by:

1. expressing the average for each school or LEA as a weighted average of the corresponding grade-level averages, where weights were determined by the number of students in each grade;
2. estimating the standard error of the mean for each grade level using standard formulas for the standard error of a mean; and
3. applying standard formulas for estimating the standard error of a weighted sum of independent random variables, as described in more detail in Appendix D.

We then used these variance estimates for the aggregate growth statistics to estimate an overall reliability for each content area, growth statistic, and level of aggregation. We provide the specific details on estimating this reliability, or ratio of the signal variance to the total variance, in Appendix D as well.

Table 11 provides the estimated reliabilities at the school- and LEA-level for each aggregated growth statistic for each content area. The reliabilities are generally all high and similar to each other, indicating that reliability, or degree of precision, does not substantially differentiate among the considered aggregated growth statistics. However, CDTM always has the lowest reliability among the three aggregated growth statistics of interest for both ELA and mathematics at both levels of aggregation.

Table 11. Reliabilities of each growth measure by content area and level of analysis

| Content Area | Growth Measure | School Level Reliability | LEA Level Reliability |
| --- | --- | --- | --- |
| ELA | CDTM | 0.87 | 0.82 |
| ELA | CPR | 0.89 | 0.86 |
| ELA | RG | 0.89 | 0.86 |
| Mathematics | CDTM | 0.91 | 0.87 |
| Mathematics | CPR | 0.92 | 0.88 |
| Mathematics | RG | 0.91 | 0.88 |

#### Interpreting Differences in Precision

There is little difference among the measures on reliability. However, CDTM’s (slightly) lower reliability for each content area and level of aggregation means that it is both sensitive to the grade composition of the school—with differential effects by content area—and does not perform better than the other measures on reliability.

# Summary of the Analyses

In this study, we evaluated three aggregated growth statistics under consideration by the California Department of Education (CDE) for measuring California Assessment of Student Performance and Progress (CAASPP) Smarter Balanced results for English language arts/literacy (ELA) and mathematics:

1. Conditional percentile rank of the gain (CPR)
2. Change in distance-to-met (CDTM)
3. Residual gain (RG)

These aggregated growth statistics were evaluated on three criteria:

1. Relationships with student demographic variables
2. Relationships with school grade composition
3. Precision/Reliability

CPR is not ideal to use for measuring growth due to its excessive relationships with aggregated student background variables. This is caused by measurement error in the prior-test score. CPR will tend to advantage schools or LEAs serving economically advantaged students even if all schools or LEAs were equally effective at promoting student growth.

CDTM, unlike the other two candidate growth statistics, is sensitive to the grade-level composition of students within a school with a differential effect by content area. Schools that serve higher grade-level students will tend to have a lower average CDTM value in ELA but a higher average CDTM value in mathematics than schools serving lower grade-level students. CDTM also has the lowest reliability across the three growth statistics, albeit not substantially lower. (When comparing reliabilities, it is important to remember that signal variance in the numerator of the reliability coefficient can be inflated by spurious relationships to demographic characteristics.)

RG more adequately addresses measurement error in the prior-test scores by using two prior-test scores (ELA and mathematics), thereby reducing the impact of student demographics on the growth results. The RG results are not highly influenced by grade configuration and RG provides a high degree of reliability. Accordingly, of the three considered aggregated student growth statistics under consideration for use in aggregate accountability, RG appears to have the most promising statistical attributes.

We understand that in addition to the technical or statistical criteria upon which we evaluated the three growth statistics, the CDE may have other practical, political, and logistical criteria upon which to evaluate the candidates. Regardless of the criteria or the prioritization of criteria used to select the aggregated student growth statistic for use in school and LEA accountability, we note that the choice of growth statistic will have practical implications for which schools and LEAs are categorized as performing well or poorly with regard to students improving over time.

The intended accountability system, which will be displayed in five-by-five colored tables, will classify schools and LEAs into one of five categories as part of the five-by-five status-by-growth accountability tables. As the particular growth statistic has not yet been selected, the thresholds for the five categories also have not been determined by the CDE. Thus, for sake of illustration, we picked quintiles (every 20th percentile; 20th, 40th, 60th, and 80th percentile) as the thresholds. We found the percentile rank of each school by each of the growth statistics, classified it into one of five categories by the quintiles, and then cross-classified with each of the other growth statistics and found the percentage of schools that would be classified into the same category (percent agreement).

Table 12 and Table 13 provide these percentage agreement statistics as well as the Spearman-rank-order correlation between each pair of growth statistics, which provides a measure of how similarly the two growth statistics rank order the schools and LEAs. Although the correlations are generally high—about .9 or above—the percentage of agreement between each pair of growth statistics tends to be moderate, with 25 to 40 percent of schools and LEAs being classified into a different category by any pair of the growth statistics. In most cases, contrasting classifications were off by just one category; however, being off by even one can be consequential: one category could be the difference between being classified as performing poorly and performing moderately. Such discrepancies are likely to occur regardless of which thresholds are chosen.

Table 12. Similarity in rank ordering and classification of schools by each pair of growth statistics

| Content Area | Score1 | Score2 | Spearman Rank Correlation | % Agreement | Total Number of Groups |
| --- | --- | --- | --- | --- | --- |
| ELA | CDTM | RG | 0.90 | 60% | 7,324 |
| ELA | CDTM | CPR | 0.89 | 58% | 7,327 |
| ELA | RG | CPR | 0.93 | 66% | 7,324 |
| Mathematics | CDTM | RG | 0.89 | 56% | 7,324 |
| Mathematics | CDTM | CPR | 0.90 | 56% | 7,325 |
| Mathematics | RG | CPR | 0.97 | 74% | 7,324 |

Table 13. Similarity in rank ordering and classification of LEAs by each pair of growth statistics

| Content Area | Score1 | Score2 | Spearman Rank Correlation | % Agreement | Total Number of Groups |
| --- | --- | --- | --- | --- | --- |
| ELA | CDTM | RG | 0.89 | 59% | 828 |
| ELA | CDTM | CPR | 0.87 | 57% | 828 |
| ELA | RG | CPR | 0.91 | 65% | 828 |
| Mathematics | CDTM | RG | 0.96 | 74% | 828 |
| Mathematics | CDTM | CPR | 0.97 | 81% | 828 |
| Mathematics | RG | CPR | 0.95 | 73% | 828 |

The extent any two growth statistics rank order and classify groups similarly varies by content area and analysis level. For all but mathematics at the LEA level, CPR and RG rank order and classify groups most similarly. In contrast, CDTM and CPR tend to rank order and classify groups the least similarly, except for mathematics at the LEA level. Thus, the extent that the choice between two measures matters in terms of real implications for the schools and LEAs depends, in part, on the content area and level of analysis. Appendix E provides the Spearman-rank-order correlations between each pair of measures aggregated at the student group level within schools and LEAs (e.g., How similar are the schools and LEAs ranked by their mean CDTM and CPR for females?). These correlations also are generally high, with variation in the most comparable measures by content area and level of analysis.

## Conclusion

This study evaluated three growth measures of interest to the CDE: change in distance-to-met, conditional percentile rank of a gain, and residual gain (using two prior scores). Among these three measures, we recommend the residual gain measure.

## Considerations for Future Research

Additional research to investigate measures that may have more desirable statistical properties is possible; three such considerations follow. It is understood that in addition to statistical considerations, there may also be practical, political, and logistical considerations that factor into the selection of the best growth measure.

First, if reporting on a percentile rank scale is of more interest for interpretability, it would be possible to take percentiles of RG for individual students prior to aggregating. These would remain about equally reliable to RG and substantially less correlated with demographic variables than CPR.

Second, RG could be improved by adjusting for measurement error in the prior scores. As discussed in the “Relationships with Student Demographic Variables” section, RG1adj should do a better job than RG of addressing measurement error in the prior score, and the empirical results suggest that this actually occurs.

Third, note that in Equation (3), the conditional expectations on the left side are assumed to have the linear regression form on the right side, but there is no guarantee such an assumption is appropriate for the data for all grades and content areas. More flexible regressions could be used and testing of the form for the prediction model might be useful. However, although rigorous testing of the suitability of the linear prediction model, and comparing alternative functional forms, was beyond the scope of this study, limited exploratory analyses of the data that we conducted did not find evidence of strong nonlinearities in the data.

# References

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Educational Measurement: Issues and Practice, 28*(4), 42–51. doi: 10.1111/j.1745-3992.2009.00161.x

Culpepper, S. A., & Aguinis, H. (2011). Using analysis of covariance (ANCOVA) with fallible covariates. *Psychological Methods*, *16*(2), 166–178. doi: 10.1037/a0023355

Fuller, W. (2006). *Measurement error models* (2nd ed.). New York: John Wiley & Sons**.**

Lockwood, J.R., & McCaffrey, D.F. (2014). Correcting for test score measurement error in ANCOVA models for estimating treatment effects. *Journal of Educational and Behavioral Measurement,* *39*(1), 22–52, 2014. doi:10.3102/1076998613509405

McCaffrey, D. F., Castellano, K. E., & Lockwood, J. R. (2015). The impactof measurement error on the accuracy of individual and aggregate SGP. *Educational Measurement: Issues and Practice*, *34*(1), 15–21. doi:10.1111/emip.12062

# Appendix A: Student Group Distributions

Table A.1. Numbers of schools and LEAs with at least 10 students in each student group

| Category | Student Group | N Schools | N LEAs |
| --- | --- | --- | --- |
| Gender | Female | 7,214 | 806 |
| Gender | Male | 7,205 | 807 |
| Race/Ethnicity | American Indian | 138 | 254 |
| Race/Ethnicity | Asian | 2,967 | 409 |
| Race/Ethnicity | Pacific Islander | 154 | 195 |
| Race/Ethnicity | Filipino | 1,258 | 315 |
| Race/Ethnicity | Hispanic | 6,919 | 748 |
| Race/Ethnicity | Black | 2,569 | 365 |
| Race/Ethnicity | White | 5,119 | 757 |
| Race/Ethnicity | Two or more races | 2,061 | 446 |
| Race/Ethnicity | Missing (Unknown) | 74 | 70 |
| English learner status | Ever EL | 6,475 | 675 |
| English learner status | Current EL | 5,702 | 625 |
| English learner status | Former EL | 5,716 | 595 |
| English learner status | Never EL | 7,277 | 822 |
| Disability status | SWD | 6,479 | 690 |
| Disability status | Not SWD | 7,311 | 825 |
| Economically disadvantaged status | SED | 7,029 | 790 |
| Economically disadvantaged status | Not SED | 6,209 | 756 |
| Homeless status | Homeless | 1,640 | 352 |
| Homeless status | Not Homeless | 7,322 | 828 |
| Foster care status | Foster youth | 2 | 112 |
| Foster care status | Not foster youth | 7,329 | 828 |

Note: These N counts of schools and LEAs are used for each subsequent graph in this appendix.

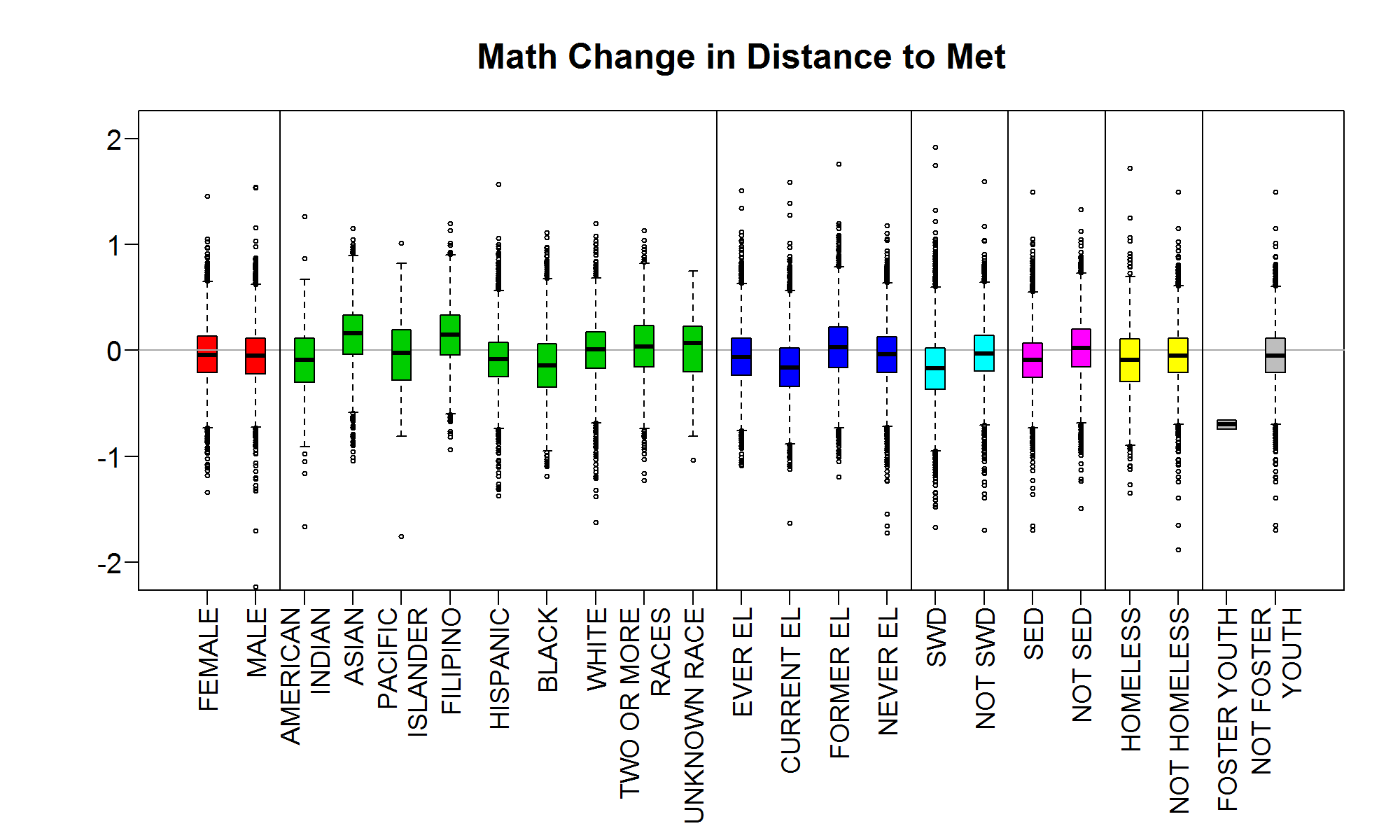


Figure A.1.A. Distributions of student group means for Change in Distance to Met in mathematics at the school level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.2. Data Supporting Figure A.1.A: Distributions of student group means for Change in Distance to Met in mathematics at the school level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.34 | -0.73 | -0.21 | -0.04 | 0.14 | 0.65 | 1.46 |
| Male | -2.23 | -0.73 | -0.22 | -0.05 | 0.12 | 0.63 | 1.55 |
| American Indian | -1.67 | -0.91 | -0.30 | -0.09 | 0.12 | 0.67 | 1.27 |
| Asian | -1.04 | -0.59 | -0.03 | 0.16 | 0.34 | 0.90 | 1.15 |
| Pacific Islander | -1.76 | -0.81 | -0.28 | -0.03 | 0.20 | 0.82 | 1.02 |
| Filipino | -0.94 | -0.60 | -0.04 | 0.15 | 0.34 | 0.90 | 1.20 |
| Hispanic | -1.37 | -0.74 | -0.25 | -0.08 | 0.08 | 0.57 | 1.57 |
| Black | -1.19 | -0.95 | -0.35 | -0.14 | 0.06 | 0.68 | 1.12 |
| White | -1.62 | -0.68 | -0.17 | 0.01 | 0.18 | 0.69 | 1.20 |
| Two or more races | -1.23 | -0.74 | -0.16 | 0.04 | 0.24 | 0.83 | 1.13 |
| Missing (Unknown) | -1.03 | -0.81 | -0.20 | 0.07 | 0.23 | 0.75 | -1.03 |
| Ever EL | -1.09 | -0.76 | -0.23 | -0.06 | 0.11 | 0.63 | 1.51 |
| Current EL | -1.63 | -0.88 | -0.34 | -0.16 | 0.02 | 0.57 | 1.59 |
| Former EL | -1.19 | -0.73 | -0.17 | 0.03 | 0.22 | 0.79 | 1.76 |
| Never EL | -1.73 | -0.72 | -0.21 | -0.04 | 0.13 | 0.64 | 1.18 |
| SWD | -3.17 | -0.95 | -0.37 | -0.17 | 0.02 | 0.60 | 1.92 |
| Not SWD | -1.70 | -0.70 | -0.20 | -0.03 | 0.14 | 0.64 | 1.60 |
| SED | -1.70 | -0.73 | -0.25 | -0.09 | 0.07 | 0.55 | 1.50 |
| Not SED | -1.49 | -0.69 | -0.15 | 0.03 | 0.20 | 0.73 | 1.33 |
| Homeless | -1.35 | -0.90 | -0.29 | -0.09 | 0.11 | 0.70 | 1.72 |
| Not Homeless | -1.89 | -0.70 | -0.21 | -0.05 | 0.12 | 0.61 | 1.50 |
| Foster youth | -0.74 | -0.74 | -0.74 | -0.70 | -0.66 | -0.66 | -0.66 |
| Not foster youth | -1.70 | -0.70 | -0.21 | -0.05 | 0.12 | 0.61 | 1.50 |

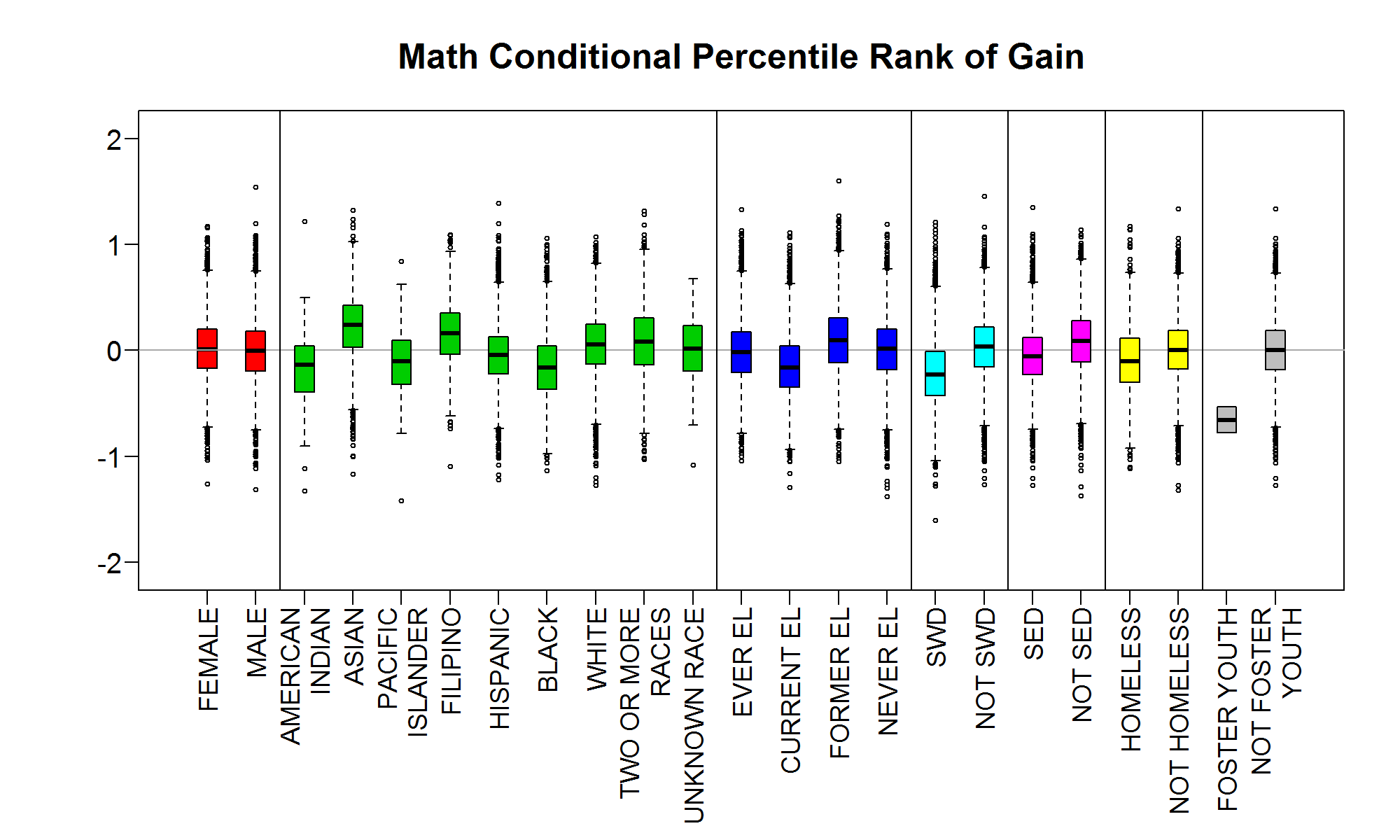


Figure A.1.B. Distributions of student group means for Conditional Percentile Rank of Gain in mathematics at the school level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.3. Data Supporting Figure A.1.B: Distributions of student group mean for Conditional Percentile Rank of Gain in mathematics at the school level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.26 | -0.73 | -0.17 | 0.01 | 0.20 | 0.76 | 1.18 |
| Male | -1.32 | -0.75 | -0.19 | 0.00 | 0.18 | 0.75 | 1.54 |
| American Indian | -1.33 | -0.90 | -0.39 | -0.14 | 0.04 | 0.50 | 1.22 |
| Asian | -1.17 | -0.56 | 0.03 | 0.24 | 0.43 | 1.03 | 1.33 |
| Pacific Islander | -1.42 | -0.79 | -0.32 | -0.10 | 0.10 | 0.62 | 0.84 |
| Filipino | -1.10 | -0.62 | -0.04 | 0.16 | 0.36 | 0.94 | 1.10 |
| Hispanic | -1.22 | -0.74 | -0.22 | -0.05 | 0.13 | 0.65 | 1.39 |
| Black | -1.14 | -0.98 | -0.37 | -0.16 | 0.04 | 0.65 | 1.06 |
| White | -1.28 | -0.70 | -0.13 | 0.06 | 0.25 | 0.82 | 1.07 |
| Two or more races | -1.03 | -0.78 | -0.13 | 0.09 | 0.31 | 0.96 | 1.32 |
| Missing (Unknown) | -1.08 | -0.70 | -0.20 | 0.02 | 0.24 | 0.68 | -1.08 |
| Ever EL | -1.04 | -0.79 | -0.21 | -0.02 | 0.18 | 0.75 | 1.33 |
| Current EL | -1.29 | -0.94 | -0.35 | -0.16 | 0.05 | 0.63 | 1.12 |
| Former EL | -1.05 | -0.74 | -0.11 | 0.09 | 0.31 | 0.94 | 1.61 |
| Never EL | -1.38 | -0.75 | -0.18 | 0.01 | 0.20 | 0.77 | 1.19 |
| SWD | -1.60 | -1.04 | -0.43 | -0.23 | -0.01 | 0.61 | 1.21 |
| Not SWD | -1.26 | -0.71 | -0.15 | 0.03 | 0.22 | 0.79 | 1.46 |
| SED | -1.27 | -0.75 | -0.23 | -0.05 | 0.12 | 0.65 | 1.35 |
| Not SED | -1.38 | -0.69 | -0.11 | 0.09 | 0.28 | 0.86 | 1.14 |
| Homeless | -1.12 | -0.92 | -0.30 | -0.11 | 0.12 | 0.74 | 1.18 |
| Not Homeless | -1.32 | -0.71 | -0.18 | 0.01 | 0.19 | 0.73 | 1.34 |
| Foster youth | -0.78 | -0.78 | -0.78 | -0.66 | -0.53 | -0.53 | -0.53 |
| Not foster youth | -1.28 | -0.73 | -0.18 | 0.00 | 0.19 | 0.73 | 1.34 |

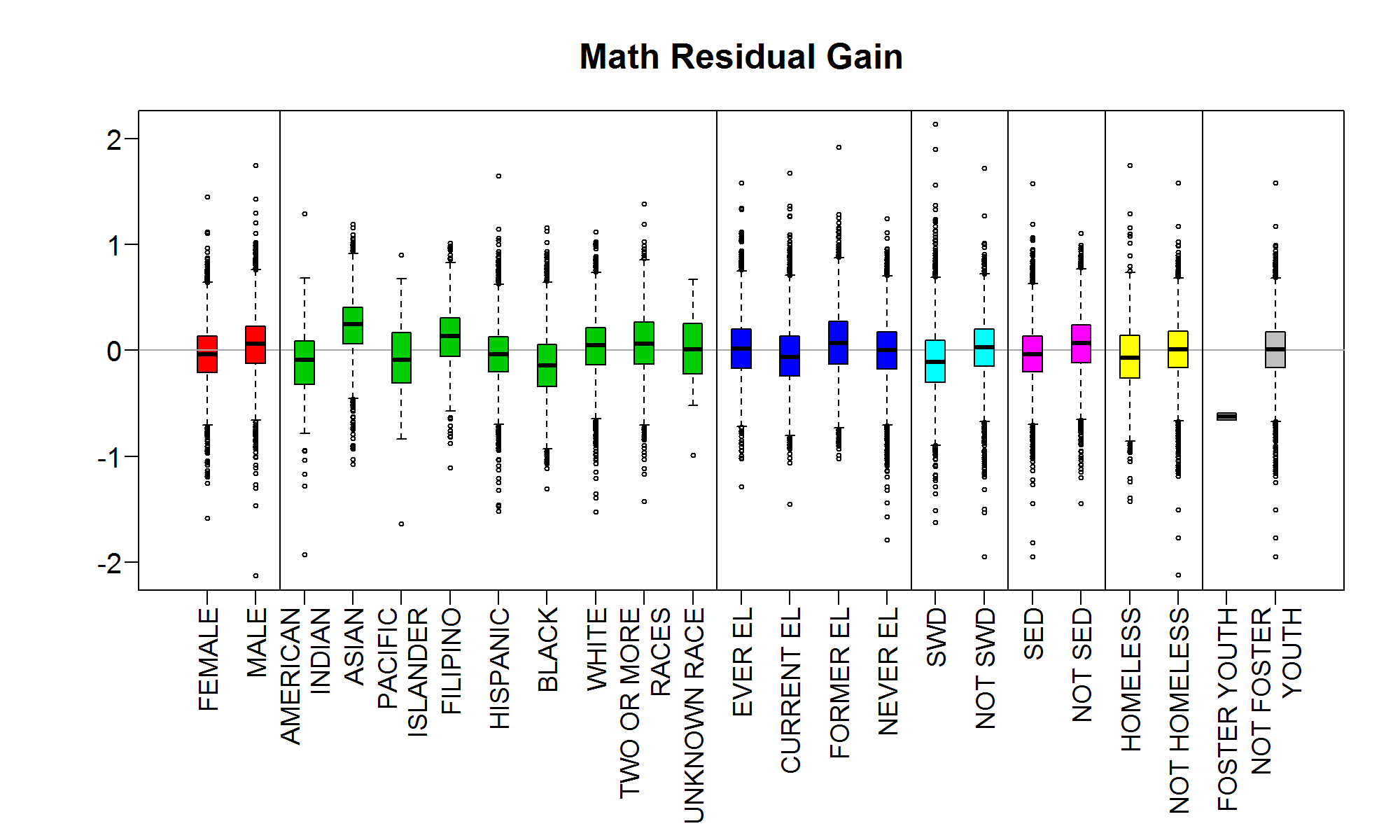


Figure A.1.C. Distributions of student group means for Residual Gain in mathematics at the school level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.4. Data Supporting Figure A.1.C: Distributions of student group means for Residual Gain in mathematics at the school level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.59 | -0.71 | -0.21 | -0.03 | 0.13 | 0.64 | 1.45 |
| Male | -2.13 | -0.66 | -0.13 | 0.06 | 0.23 | 0.76 | 1.75 |
| American Indian | -1.93 | -0.79 | -0.32 | -0.09 | 0.09 | 0.69 | 1.29 |
| Asian | -1.08 | -0.46 | 0.06 | 0.25 | 0.41 | 0.92 | 1.19 |
| Pacific Islander | -1.64 | -0.84 | -0.31 | -0.09 | 0.17 | 0.68 | 0.90 |
| Filipino | -1.11 | -0.58 | -0.05 | 0.14 | 0.31 | 0.83 | 1.01 |
| Hispanic | -1.52 | -0.70 | -0.20 | -0.04 | 0.13 | 0.62 | 1.65 |
| Black | -1.31 | -0.93 | -0.34 | -0.15 | 0.06 | 0.65 | 1.16 |
| White | -1.52 | -0.64 | -0.13 | 0.05 | 0.22 | 0.74 | 1.12 |
| Two or more races | -1.43 | -0.71 | -0.13 | 0.06 | 0.27 | 0.86 | 1.39 |
| Missing (Unknown) | -0.99 | -0.52 | -0.22 | 0.01 | 0.25 | 0.67 | -0.99 |
| Ever EL | -1.29 | -0.72 | -0.17 | 0.02 | 0.20 | 0.75 | 1.58 |
| Current EL | -1.45 | -0.81 | -0.24 | -0.07 | 0.14 | 0.71 | 1.68 |
| Former EL | -1.03 | -0.73 | -0.13 | 0.07 | 0.27 | 0.88 | 1.93 |
| Never EL | -1.79 | -0.70 | -0.18 | 0.00 | 0.18 | 0.71 | 1.25 |
| SWD | -3.17 | -0.90 | -0.30 | -0.11 | 0.10 | 0.69 | 2.14 |
| Not SWD | -1.95 | -0.67 | -0.15 | 0.03 | 0.20 | 0.72 | 1.72 |
| SED | -1.95 | -0.70 | -0.20 | -0.04 | 0.13 | 0.64 | 1.58 |
| Not SED | -1.45 | -0.65 | -0.12 | 0.07 | 0.24 | 0.77 | 1.11 |
| Homeless | -1.42 | -0.86 | -0.26 | -0.07 | 0.14 | 0.74 | 1.75 |
| Not Homeless | -2.12 | -0.67 | -0.16 | 0.01 | 0.18 | 0.69 | 1.58 |
| Foster youth | -0.66 | -0.66 | -0.66 | -0.63 | -0.59 | -0.59 | -0.59 |
| Not foster youth | -1.95 | -0.67 | -0.16 | 0.01 | 0.18 | 0.69 | 1.58 |

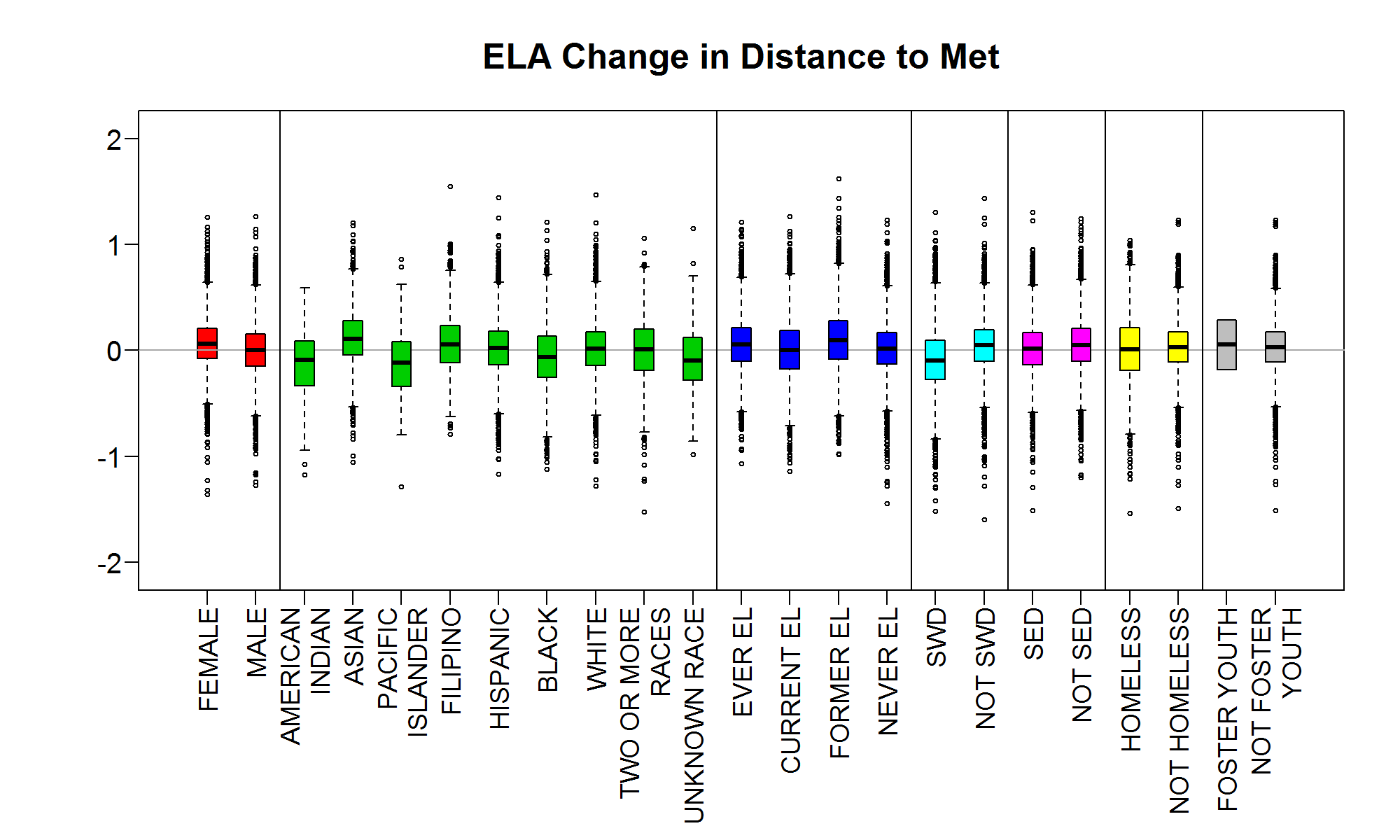


Figure A.2.A. Distributions of student group means for Change in Distance to Met in ELA at the school level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.5 Data Supporting Figure A.2.A: Distributions of student group means for Change in Distance to Met in ELA at the school level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.36 | -0.51 | -0.08 | 0.06 | 0.21 | 0.64 | 1.26 |
| Male | -1.27 | -0.62 | -0.15 | 0.00 | 0.16 | 0.62 | 1.27 |
| American Indian | -1.17 | -0.94 | -0.34 | -0.09 | 0.09 | 0.59 | -1.08 |
| Asian | -1.06 | -0.53 | -0.05 | 0.11 | 0.28 | 0.77 | 1.21 |
| Pacific Islander | -1.29 | -0.80 | -0.34 | -0.12 | 0.08 | 0.62 | 0.86 |
| Filipino | -0.79 | -0.63 | -0.11 | 0.06 | 0.24 | 0.76 | 1.55 |
| Hispanic | -1.17 | -0.60 | -0.13 | 0.02 | 0.18 | 0.65 | 1.45 |
| Black | -1.12 | -0.82 | -0.25 | -0.06 | 0.14 | 0.72 | 1.22 |
| White | -1.28 | -0.62 | -0.14 | 0.02 | 0.18 | 0.65 | 1.47 |
| Two or more races | -1.53 | -0.77 | -0.19 | 0.01 | 0.20 | 0.79 | 1.06 |
| Missing (Unknown) | -0.98 | -0.86 | -0.28 | -0.09 | 0.12 | 0.71 | 1.15 |
| Ever EL | -1.07 | -0.58 | -0.10 | 0.05 | 0.22 | 0.69 | 1.21 |
| Current EL | -1.14 | -0.71 | -0.18 | 0.01 | 0.19 | 0.73 | 1.27 |
| Former EL | -0.99 | -0.62 | -0.08 | 0.10 | 0.28 | 0.82 | 1.63 |
| Never EL | -1.45 | -0.57 | -0.13 | 0.02 | 0.17 | 0.61 | 1.23 |
| SWD | -3.48 | -0.83 | -0.28 | -0.09 | 0.09 | 0.64 | 1.31 |
| Not SWD | -1.60 | -0.54 | -0.10 | 0.05 | 0.20 | 0.64 | 1.44 |
| SED | -1.51 | -0.59 | -0.14 | 0.02 | 0.17 | 0.62 | 1.31 |
| Not SED | -1.20 | -0.57 | -0.10 | 0.05 | 0.21 | 0.67 | 1.25 |
| Homeless | -1.54 | -0.79 | -0.19 | 0.01 | 0.21 | 0.81 | 1.04 |
| Not Homeless | -1.49 | -0.54 | -0.11 | 0.03 | 0.17 | 0.60 | 1.23 |
| Foster youth | -0.18 | -0.18 | -0.18 | 0.05 | 0.29 | 0.29 | 0.29 |
| Not foster youth | -1.51 | -0.53 | -0.11 | 0.03 | 0.17 | 0.59 | 1.23 |

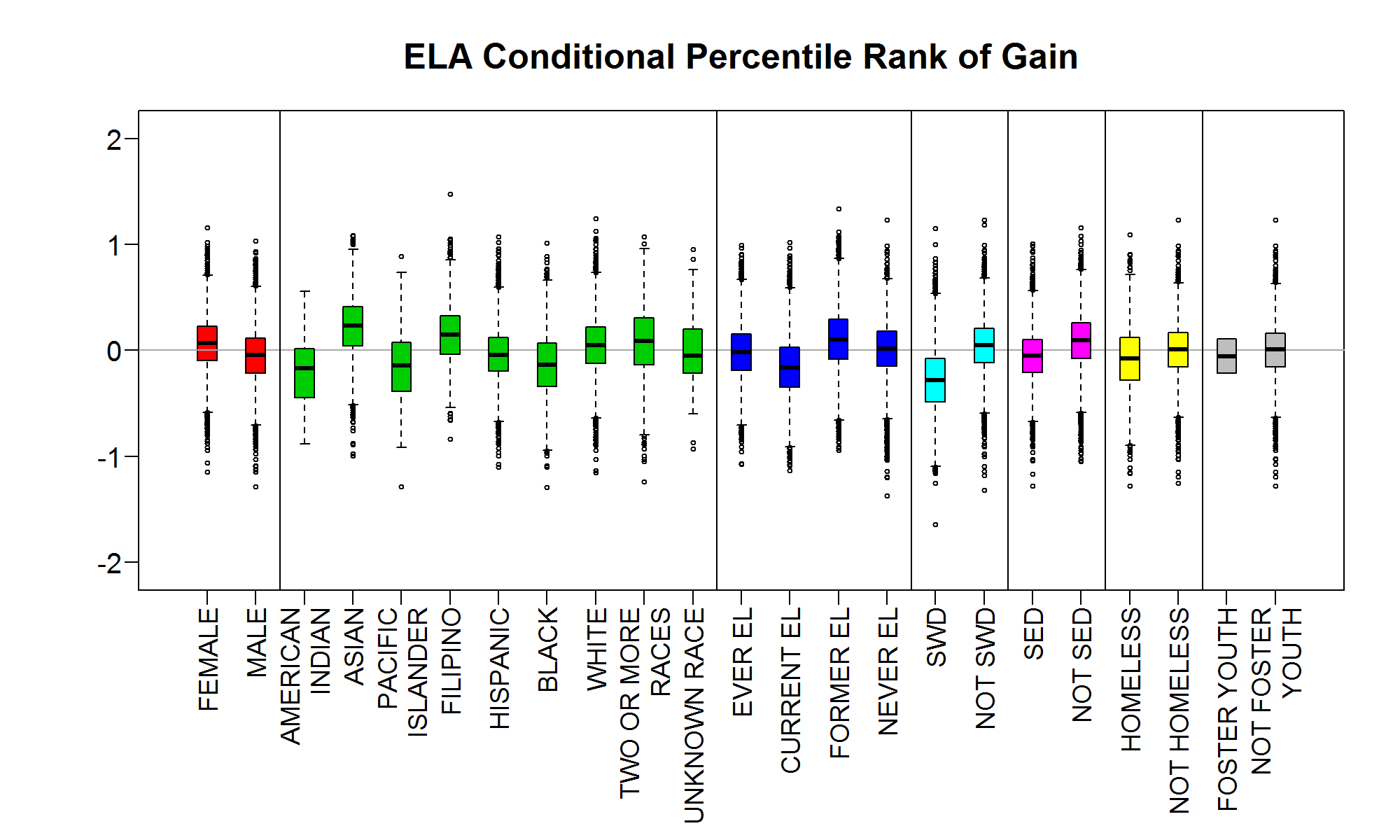


Figure A.2.B. Distributions of student group means for Conditional Percentile Rank of Gain in ELA at the school level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.6. Data Supporting Figure A.2.B: Distributions of student group means for Conditional Percentile Rank of Gain in ELA at the school level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.15 | -0.58 | -0.10 | 0.07 | 0.23 | 0.71 | 1.16 |
| Male | -1.29 | -0.70 | -0.22 | -0.05 | 0.12 | 0.61 | 1.04 |
| American Indian | -0.89 | -0.89 | -0.45 | -0.17 | 0.02 | 0.56 | 0.56 |
| Asian | -1.00 | -0.52 | 0.04 | 0.24 | 0.41 | 0.95 | 1.09 |
| Pacific Islander | -1.29 | -0.91 | -0.39 | -0.14 | 0.08 | 0.74 | 0.89 |
| Filipino | -0.84 | -0.54 | -0.04 | 0.15 | 0.33 | 0.86 | 1.48 |
| Hispanic | -1.11 | -0.68 | -0.20 | -0.04 | 0.12 | 0.60 | 1.08 |
| Black | -1.30 | -0.95 | -0.34 | -0.14 | 0.07 | 0.67 | 1.02 |
| White | -1.15 | -0.64 | -0.12 | 0.05 | 0.22 | 0.74 | 1.25 |
| Two or more races | -1.24 | -0.80 | -0.14 | 0.09 | 0.31 | 0.96 | 1.08 |
| Missing (Unknown) | -0.93 | -0.60 | -0.22 | -0.05 | 0.20 | 0.76 | 0.95 |
| Ever EL | -1.08 | -0.71 | -0.19 | -0.02 | 0.16 | 0.67 | 1.00 |
| Current EL | -1.14 | -0.91 | -0.35 | -0.16 | 0.03 | 0.59 | 1.02 |
| Former EL | -0.94 | -0.66 | -0.09 | 0.11 | 0.30 | 0.87 | 1.34 |
| Never EL | -1.37 | -0.65 | -0.15 | 0.02 | 0.18 | 0.68 | 1.23 |
| SWD | -1.65 | -1.10 | -0.48 | -0.28 | -0.08 | 0.54 | 1.15 |
| Not SWD | -1.32 | -0.59 | -0.11 | 0.05 | 0.21 | 0.69 | 1.23 |
| SED | -1.28 | -0.67 | -0.21 | -0.05 | 0.10 | 0.57 | 1.01 |
| Not SED | -1.05 | -0.58 | -0.08 | 0.09 | 0.26 | 0.77 | 1.16 |
| Homeless | -1.28 | -0.89 | -0.28 | -0.08 | 0.12 | 0.72 | 1.10 |
| Not Homeless | -1.25 | -0.63 | -0.15 | 0.01 | 0.17 | 0.64 | 1.23 |
| Foster youth | -0.22 | -0.22 | -0.22 | -0.05 | 0.11 | 0.11 | 0.11 |
| Not foster youth | -1.28 | -0.63 | -0.15 | 0.01 | 0.16 | 0.63 | 1.23 |

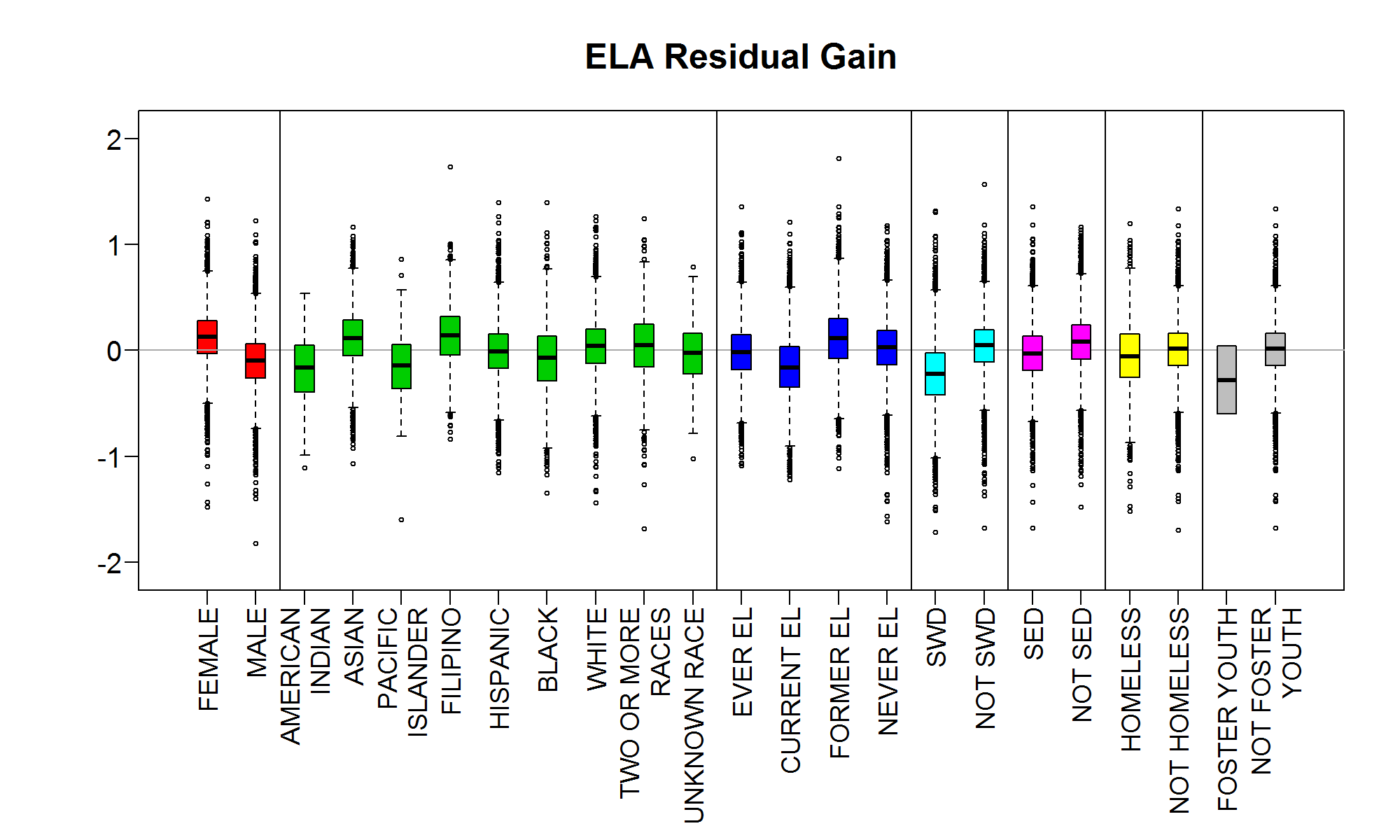


Figure A.2.C. Distributions of student group means for Residual Gain in ELA at the school level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.7 Data Supporting Figure A.2.C: Distributions of student group means for Residual Gain in ELA at the school level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.48 | -0.50 | -0.03 | 0.13 | 0.28 | 0.75 | 1.43 |
| Male | -1.82 | -0.74 | -0.26 | -0.10 | 0.06 | 0.54 | 1.22 |
| American Indian | -1.11 | -0.99 | -0.39 | -0.16 | 0.05 | 0.54 | -1.11 |
| Asian | -1.07 | -0.54 | -0.05 | 0.11 | 0.28 | 0.78 | 1.17 |
| Pacific Islander | -1.60 | -0.81 | -0.36 | -0.14 | 0.06 | 0.57 | 0.86 |
| Filipino | -0.83 | -0.59 | -0.04 | 0.14 | 0.32 | 0.86 | 1.74 |
| Hispanic | -1.16 | -0.66 | -0.17 | -0.01 | 0.15 | 0.64 | 1.40 |
| Black | -1.35 | -0.93 | -0.29 | -0.07 | 0.14 | 0.77 | 1.40 |
| White | -1.44 | -0.62 | -0.12 | 0.04 | 0.20 | 0.70 | 1.27 |
| Two or more races | -1.69 | -0.75 | -0.16 | 0.05 | 0.25 | 0.84 | 1.25 |
| Missing (Unknown) | -1.02 | -0.79 | -0.22 | -0.02 | 0.16 | 0.70 | 0.79 |
| Ever EL | -1.09 | -0.68 | -0.18 | -0.02 | 0.15 | 0.65 | 1.36 |
| Current EL | -1.22 | -0.90 | -0.35 | -0.16 | 0.03 | 0.60 | 1.21 |
| Former EL | -1.12 | -0.65 | -0.08 | 0.11 | 0.30 | 0.87 | 1.81 |
| Never EL | -1.62 | -0.61 | -0.13 | 0.03 | 0.19 | 0.66 | 1.18 |
| SWD | -3.95 | -1.01 | -0.42 | -0.22 | -0.02 | 0.57 | 1.32 |
| Not SWD | -1.68 | -0.57 | -0.11 | 0.05 | 0.20 | 0.65 | 1.57 |
| SED | -1.68 | -0.67 | -0.19 | -0.03 | 0.13 | 0.61 | 1.36 |
| Not SED | -1.48 | -0.57 | -0.08 | 0.08 | 0.24 | 0.73 | 1.17 |
| Homeless | -1.52 | -0.87 | -0.26 | -0.05 | 0.16 | 0.78 | 1.20 |
| Not Homeless | -1.70 | -0.59 | -0.14 | 0.01 | 0.16 | 0.61 | 1.34 |
| Foster youth | -0.60 | -0.60 | -0.60 | -0.28 | 0.04 | 0.04 | 0.04 |
| Not foster youth | -1.68 | -0.60 | -0.14 | 0.01 | 0.16 | 0.61 | 1.34 |

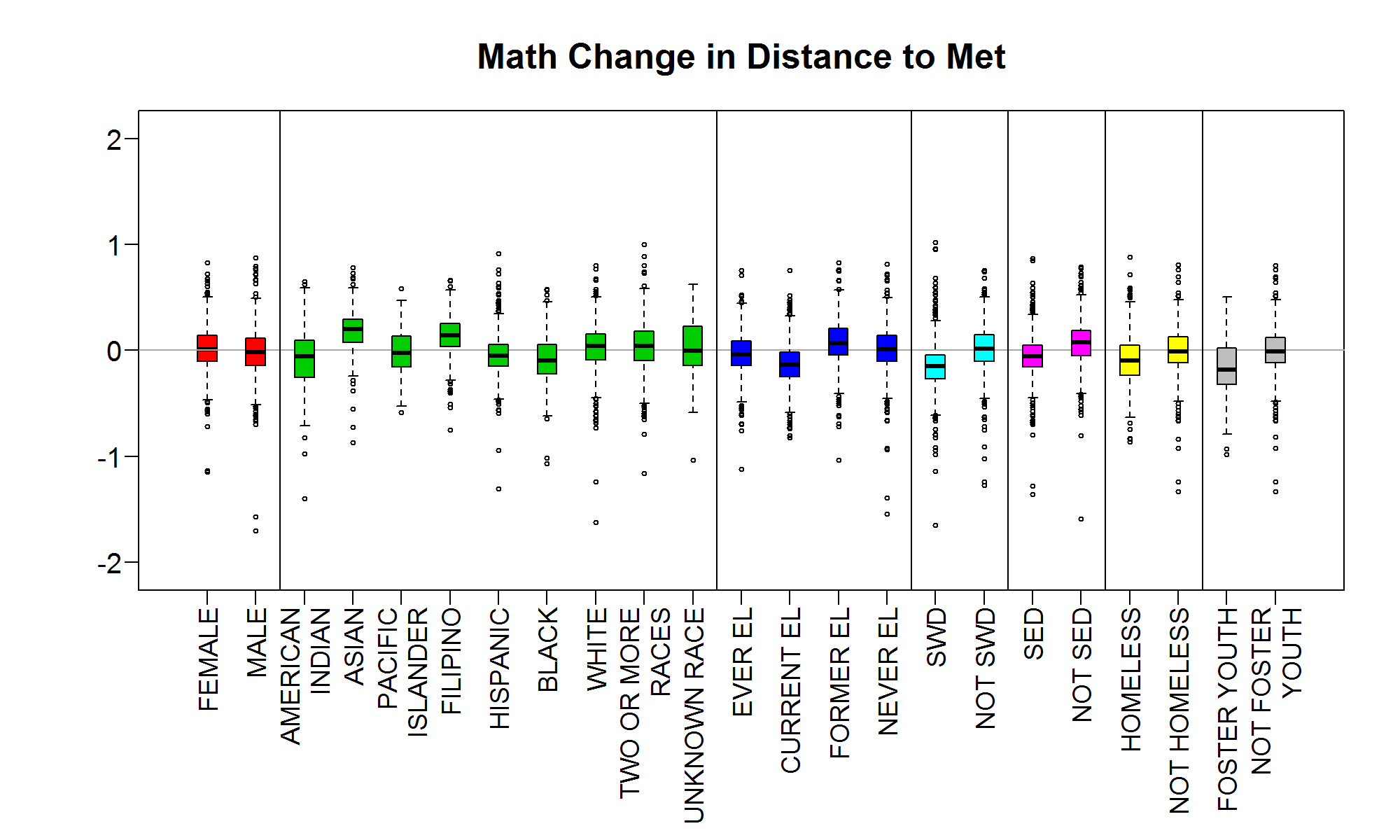


Figure A.3.A. Distributions of student group means for Change in Distance to Met for mathematics at the LEA level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.8. Data Supporting Figure A.3.A: Distributions of student group means for Change in Distance to Met in mathematics at the LEA level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.15 | -0.47 | -0.10 | 0.01 | 0.15 | 0.51 | 0.83 |
| Male | -1.70 | -0.52 | -0.14 | -0.02 | 0.12 | 0.49 | 0.88 |
| American Indian | -1.40 | -0.71 | -0.25 | -0.06 | 0.10 | 0.59 | 0.65 |
| Asian | -0.87 | -0.24 | 0.08 | 0.20 | 0.29 | 0.59 | 0.78 |
| Pacific Islander | -0.59 | -0.53 | -0.16 | -0.02 | 0.13 | 0.47 | 0.58 |
| Filipino | -0.75 | -0.28 | 0.04 | 0.14 | 0.25 | 0.57 | 0.66 |
| Hispanic | -1.31 | -0.46 | -0.15 | -0.05 | 0.06 | 0.35 | 0.92 |
| Black | -1.07 | -0.62 | -0.22 | -0.09 | 0.05 | 0.46 | 0.58 |
| White | -1.62 | -0.45 | -0.09 | 0.04 | 0.15 | 0.51 | 0.80 |
| Two or more races | -1.16 | -0.50 | -0.09 | 0.04 | 0.18 | 0.58 | 1.00 |
| Missing (Unknown) | -1.03 | -0.59 | -0.14 | 0.00 | 0.23 | 0.63 | -1.03 |
| Ever EL | -1.12 | -0.49 | -0.15 | -0.04 | 0.09 | 0.44 | 0.76 |
| Current EL | -0.82 | -0.58 | -0.25 | -0.14 | -0.02 | 0.33 | 0.76 |
| Former EL | -1.04 | -0.41 | -0.04 | 0.07 | 0.21 | 0.57 | 0.83 |
| Never EL | -1.54 | -0.45 | -0.10 | 0.01 | 0.14 | 0.50 | 0.82 |
| SWD | -1.65 | -0.61 | -0.27 | -0.15 | -0.04 | 0.28 | 1.03 |
| Not SWD | -1.28 | -0.46 | -0.10 | 0.02 | 0.15 | 0.51 | 0.76 |
| SED | -1.36 | -0.45 | -0.15 | -0.06 | 0.05 | 0.34 | 0.87 |
| Not SED | -1.59 | -0.40 | -0.05 | 0.08 | 0.19 | 0.52 | 0.79 |
| Homeless | -0.86 | -0.63 | -0.24 | -0.10 | 0.05 | 0.46 | 0.88 |
| Not Homeless | -1.33 | -0.48 | -0.12 | -0.01 | 0.13 | 0.48 | 0.81 |
| Foster youth | -0.98 | -0.79 | -0.32 | -0.18 | 0.02 | 0.50 | -0.93 |
| Not foster youth | -1.33 | -0.48 | -0.12 | -0.01 | 0.13 | 0.48 | 0.81 |

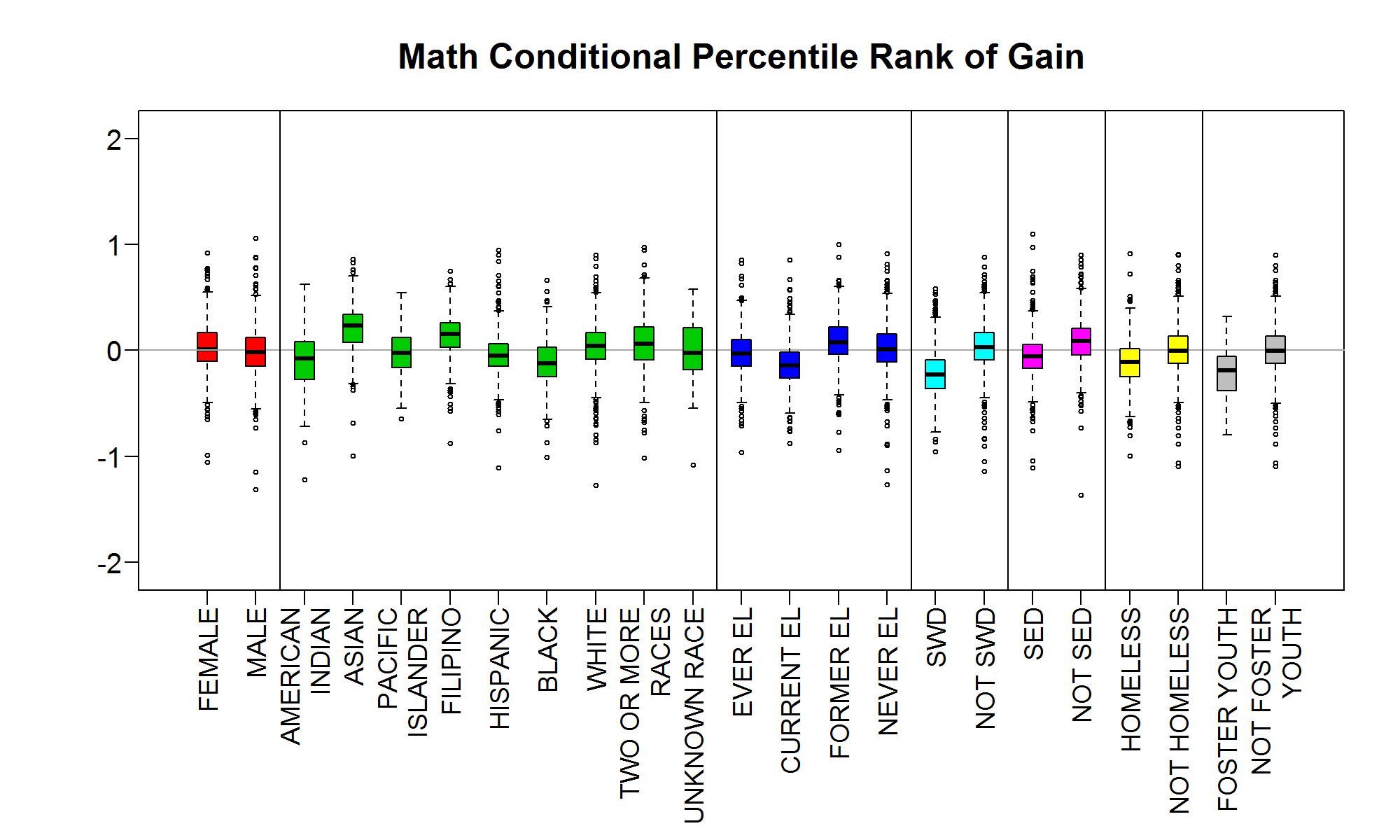


Figure A.3.B. Distributions of student group means for Conditional Percentile Rank of Gain in mathematics at the LEA level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.9. Data Supporting Figure A.3.B: Distributions of student group means for Conditional Percentile Rank of Gain in mathematics at the LEA level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.06 | -0.49 | -0.10 | 0.01 | 0.17 | 0.55 | 0.92 |
| Male | -1.32 | -0.55 | -0.15 | -0.01 | 0.12 | 0.52 | 1.06 |
| American Indian | -1.22 | -0.72 | -0.28 | -0.08 | 0.09 | 0.62 | -0.87 |
| Asian | -0.99 | -0.31 | 0.08 | 0.23 | 0.34 | 0.70 | 0.86 |
| Pacific Islander | -0.65 | -0.55 | -0.16 | -0.02 | 0.13 | 0.55 | -0.65 |
| Filipino | -0.88 | -0.32 | 0.03 | 0.16 | 0.26 | 0.60 | 0.75 |
| Hispanic | -1.11 | -0.47 | -0.15 | -0.05 | 0.06 | 0.37 | 0.95 |
| Black | -1.01 | -0.65 | -0.25 | -0.12 | 0.03 | 0.41 | 0.67 |
| White | -1.28 | -0.45 | -0.08 | 0.04 | 0.17 | 0.54 | 0.90 |
| Two or more races | -1.01 | -0.50 | -0.09 | 0.06 | 0.22 | 0.68 | 0.98 |
| Missing (Unknown) | -1.08 | -0.55 | -0.18 | -0.02 | 0.22 | 0.58 | -1.08 |
| Ever EL | -0.97 | -0.50 | -0.15 | -0.03 | 0.10 | 0.47 | 0.86 |
| Current EL | -0.87 | -0.59 | -0.26 | -0.14 | -0.02 | 0.34 | 0.86 |
| Former EL | -0.94 | -0.42 | -0.04 | 0.08 | 0.22 | 0.60 | 1.00 |
| Never EL | -1.27 | -0.47 | -0.11 | 0.01 | 0.15 | 0.54 | 0.92 |
| SWD | -0.95 | -0.77 | -0.36 | -0.23 | -0.09 | 0.31 | 0.59 |
| Not SWD | -1.14 | -0.44 | -0.09 | 0.03 | 0.17 | 0.55 | 0.89 |
| SED | -1.11 | -0.48 | -0.17 | -0.06 | 0.05 | 0.37 | 1.10 |
| Not SED | -1.37 | -0.40 | -0.04 | 0.09 | 0.21 | 0.58 | 0.90 |
| Homeless | -0.99 | -0.63 | -0.25 | -0.11 | 0.02 | 0.40 | 0.92 |
| Not Homeless | -1.10 | -0.49 | -0.12 | 0.00 | 0.14 | 0.52 | 0.91 |
| Foster youth | -0.80 | -0.80 | -0.38 | -0.19 | -0.06 | 0.32 | 0.32 |
| Not foster youth | -1.10 | -0.50 | -0.12 | 0.00 | 0.14 | 0.52 | 0.91 |

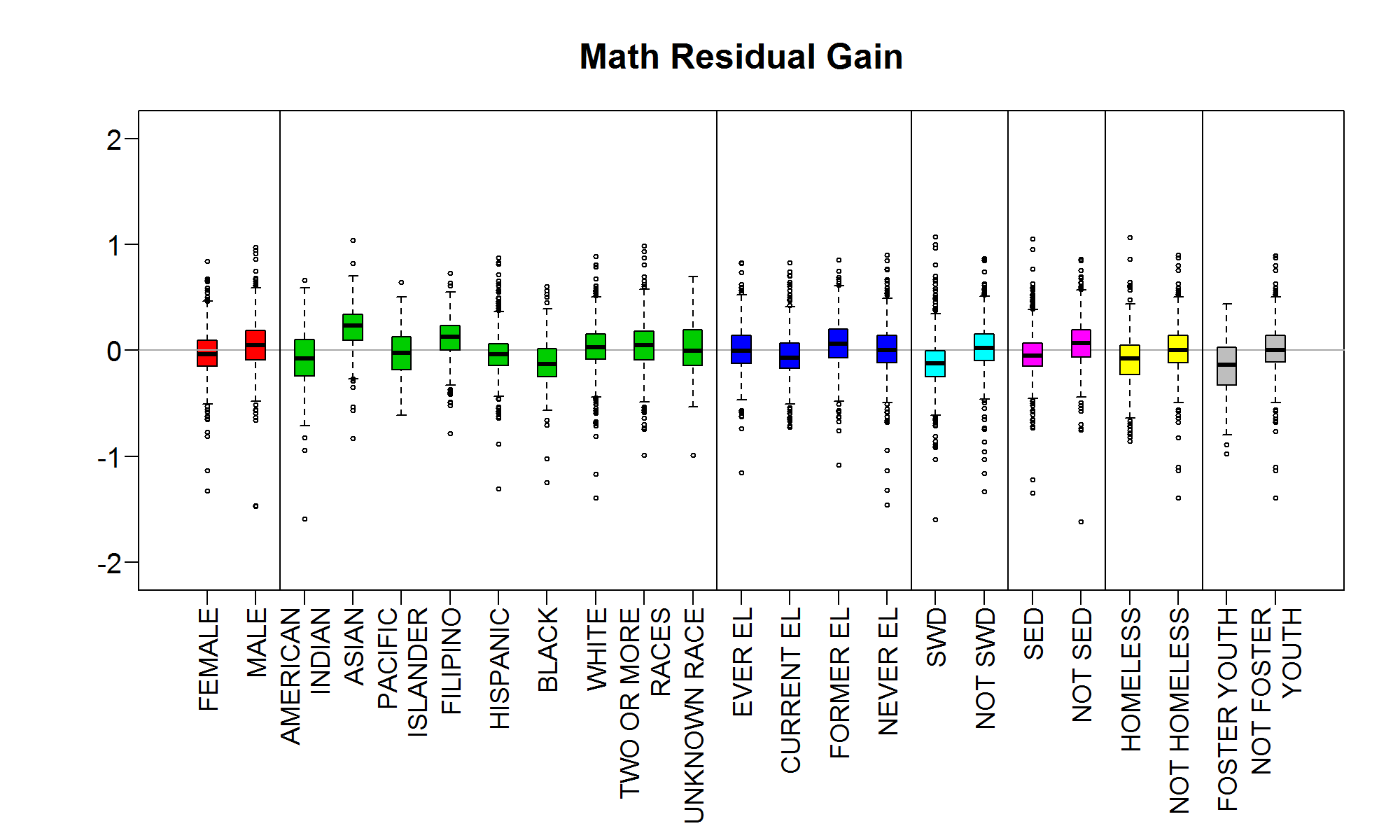


Figure A.3.C. Distributions of student group growth means for Residual Gain in mathematics at the LEA level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.10. Data Supporting Figure A.3.C: Distributions of student group means for Residual Gain in mathematics at the LEA level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.33 | -0.51 | -0.15 | -0.04 | 0.10 | 0.46 | 0.84 |
| Male | -1.47 | -0.48 | -0.09 | 0.05 | 0.19 | 0.59 | 0.97 |
| American Indian | -1.59 | -0.71 | -0.24 | -0.08 | 0.10 | 0.59 | 0.66 |
| Asian | -0.83 | -0.27 | 0.10 | 0.24 | 0.34 | 0.71 | 1.04 |
| Pacific Islander | 0.64 | -0.61 | -0.18 | -0.02 | 0.13 | 0.51 | 0.64 |
| Filipino | -0.78 | -0.33 | 0.01 | 0.13 | 0.24 | 0.55 | 0.73 |
| Hispanic | -1.31 | -0.43 | -0.14 | -0.04 | 0.06 | 0.37 | 0.88 |
| Black | -1.25 | -0.57 | -0.25 | -0.13 | 0.01 | 0.40 | 0.60 |
| White | -1.39 | -0.44 | -0.08 | 0.03 | 0.16 | 0.51 | 0.89 |
| Two or more races | -0.99 | -0.49 | -0.09 | 0.05 | 0.18 | 0.58 | 0.99 |
| Missing (Unknown) | -0.99 | -0.53 | -0.14 | 0.00 | 0.20 | 0.70 | -0.99 |
| Ever EL | -1.16 | -0.46 | -0.12 | 0.00 | 0.14 | 0.52 | 0.83 |
| Current EL | -0.73 | -0.51 | -0.17 | -0.07 | 0.07 | 0.41 | 0.83 |
| Former EL | -1.08 | -0.48 | -0.07 | 0.06 | 0.20 | 0.61 | 0.85 |
| Never EL | -1.46 | -0.49 | -0.11 | 0.00 | 0.14 | 0.49 | 0.90 |
| SWD | -1.60 | -0.61 | -0.25 | -0.13 | 0.00 | 0.35 | 1.08 |
| Not SWD | -1.34 | -0.46 | -0.09 | 0.02 | 0.15 | 0.51 | 0.87 |
| SED | -1.35 | -0.45 | -0.15 | -0.05 | 0.07 | 0.38 | 1.06 |
| Not SED | -1.62 | -0.44 | -0.06 | 0.07 | 0.19 | 0.58 | 0.86 |
| Homeless | -0.86 | -0.64 | -0.23 | -0.08 | 0.05 | 0.44 | 1.07 |
| Not Homeless | -1.39 | -0.49 | -0.11 | 0.01 | 0.14 | 0.51 | 0.90 |
| Foster youth | -0.98 | -0.80 | -0.33 | -0.14 | 0.03 | 0.44 | -0.89 |
| Not foster youth | -1.39 | -0.49 | -0.11 | 0.00 | 0.14 | 0.51 | 0.89 |

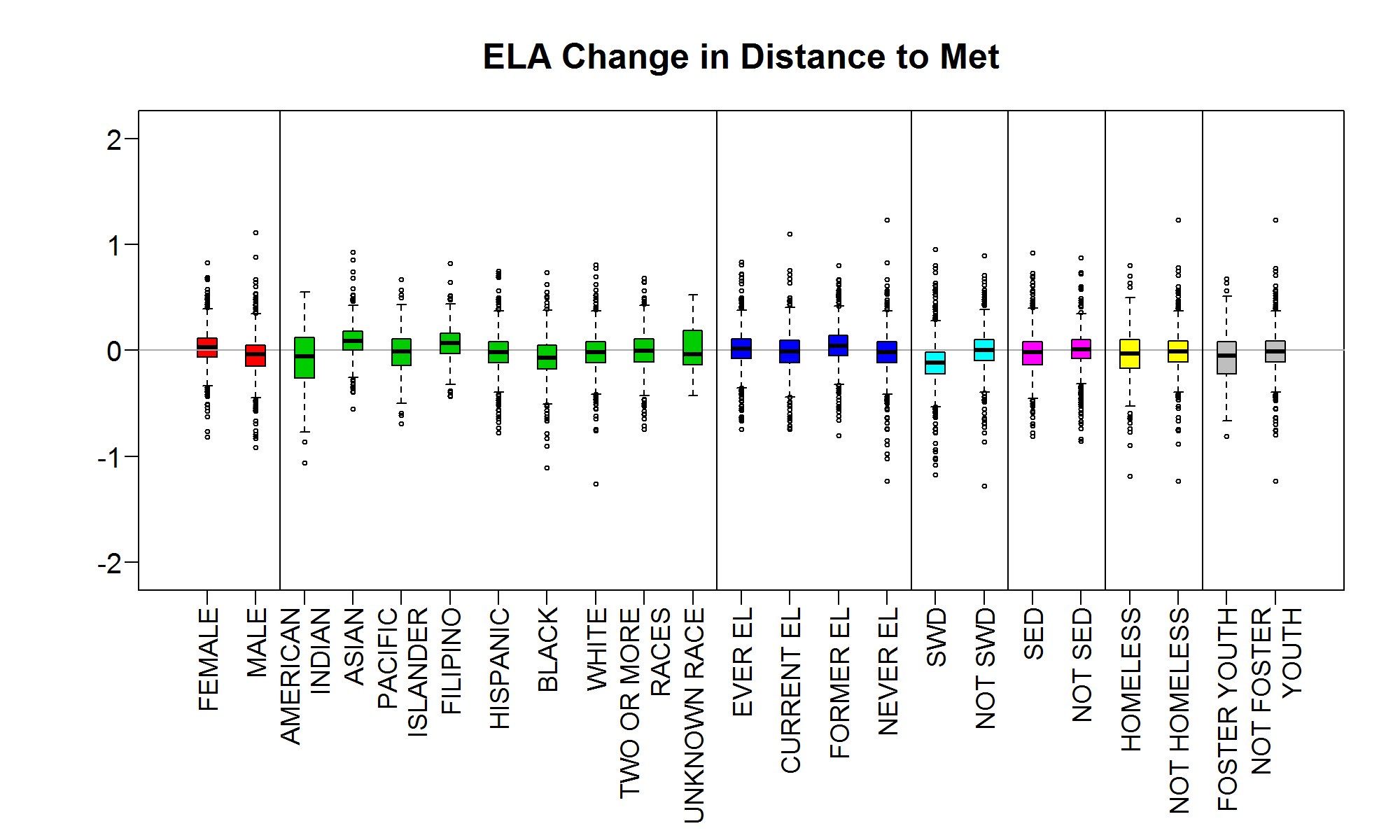


Figure A.4.A. Distributions of student group means for Change in Distance to Met in ELA at the LEA level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.11. Data Supporting Figure A.4.A: Distributions of student group means for Change in Distance to Met in ELA at the LEA level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -0.82 | -0.33 | -0.06 | 0.03 | 0.12 | 0.39 | 0.83 |
| Male | -0.91 | -0.45 | -0.15 | -0.04 | 0.05 | 0.35 | 1.12 |
| American Indian | -1.06 | -0.77 | -0.26 | -0.06 | 0.12 | 0.55 | -0.86 |
| Asian | -0.55 | -0.25 | 0.00 | 0.09 | 0.18 | 0.43 | 0.93 |
| Pacific Islander | -0.69 | -0.50 | -0.14 | -0.01 | 0.11 | 0.44 | 0.67 |
| Filipino | -0.43 | -0.32 | -0.03 | 0.07 | 0.16 | 0.44 | 0.82 |
| Hispanic | -0.78 | -0.39 | -0.11 | -0.02 | 0.08 | 0.38 | 0.75 |
| Black | -1.11 | -0.51 | -0.18 | -0.07 | 0.05 | 0.38 | 0.74 |
| White | -1.26 | -0.41 | -0.12 | -0.01 | 0.08 | 0.37 | 0.81 |
| Two or more races | -0.74 | -0.43 | -0.11 | 0.00 | 0.11 | 0.43 | 0.69 |
| Missing (Unknown) | -0.43 | -0.43 | -0.13 | -0.03 | 0.19 | 0.53 | 0.53 |
| Ever EL | -0.74 | -0.35 | -0.08 | 0.02 | 0.11 | 0.38 | 0.84 |
| Current EL | -0.74 | -0.44 | -0.12 | -0.01 | 0.10 | 0.41 | 1.10 |
| Former EL | -0.80 | -0.32 | -0.05 | 0.04 | 0.14 | 0.42 | 0.81 |
| Never EL | -1.23 | -0.41 | -0.12 | -0.02 | 0.08 | 0.37 | 1.23 |
| SWD | -1.18 | -0.53 | -0.22 | -0.11 | -0.02 | 0.28 | 0.96 |
| Not SWD | -1.28 | -0.39 | -0.10 | 0.01 | 0.10 | 0.38 | 0.89 |
| SED | -0.81 | -0.45 | -0.13 | -0.02 | 0.08 | 0.40 | 0.92 |
| Not SED | -0.86 | -0.32 | -0.07 | 0.01 | 0.10 | 0.34 | 0.87 |
| Homeless | -1.19 | -0.52 | -0.17 | -0.03 | 0.11 | 0.50 | 0.80 |
| Not Homeless | -1.23 | -0.39 | -0.11 | -0.01 | 0.09 | 0.37 | 1.23 |
| Foster youth | -0.81 | -0.67 | -0.22 | -0.05 | 0.08 | 0.51 | 0.68 |
| Not foster youth | -1.23 | -0.40 | -0.11 | -0.01 | 0.09 | 0.37 | 1.23 |

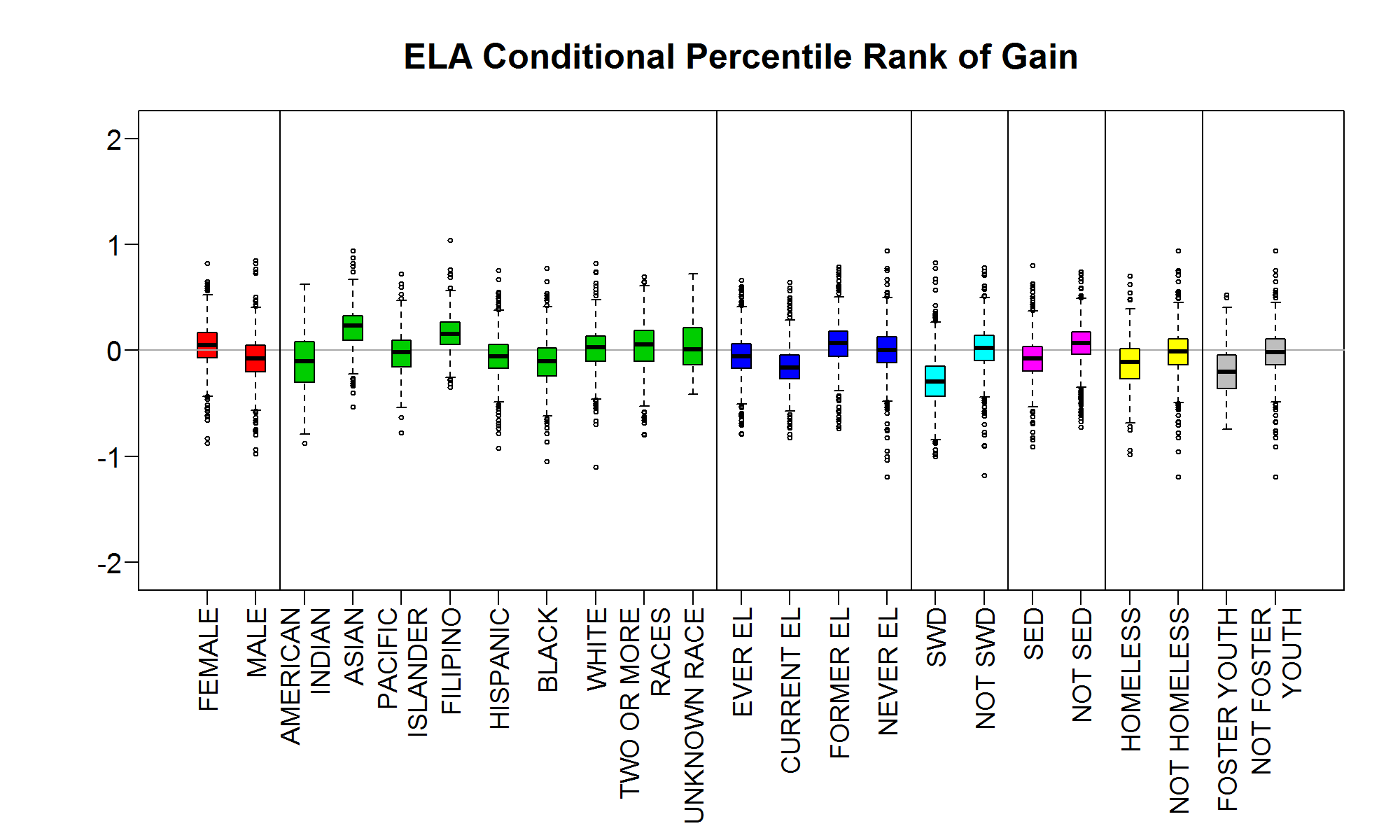


Figure A.4.B. Distributions of student group means for Conditional Percentile Rank of Gain in ELA at the LEA level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.12. Data Supporting Figure A.4.B: Distributions of student group means for Conditional Percentile Rank of Gain in ELA at the LEA level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -0.88 | -0.43 | -0.07 | 0.05 | 0.17 | 0.53 | 0.82 |
| Male | -0.98 | -0.57 | -0.20 | -0.07 | 0.05 | 0.41 | 0.85 |
| American Indian | -0.88 | -0.79 | -0.30 | -0.10 | 0.08 | 0.62 | -0.88 |
| Asian | -0.53 | -0.22 | 0.10 | 0.23 | 0.33 | 0.67 | 0.94 |
| Pacific Islander | -0.77 | -0.54 | -0.16 | -0.02 | 0.10 | 0.47 | 0.72 |
| Filipino | -0.35 | -0.26 | 0.06 | 0.15 | 0.27 | 0.56 | 1.04 |
| Hispanic | -0.92 | -0.49 | -0.17 | -0.06 | 0.05 | 0.38 | 0.76 |
| Black | -1.05 | -0.62 | -0.24 | -0.10 | 0.02 | 0.42 | 0.78 |
| White | -1.10 | -0.46 | -0.10 | 0.03 | 0.13 | 0.48 | 0.82 |
| Two or more races | -0.80 | -0.53 | -0.10 | 0.06 | 0.19 | 0.62 | 0.70 |
| Missing (Unknown) | -0.42 | -0.42 | -0.14 | 0.01 | 0.22 | 0.72 | 0.72 |
| Ever EL | -0.79 | -0.51 | -0.17 | -0.06 | 0.07 | 0.41 | 0.66 |
| Current EL | -0.82 | -0.57 | -0.27 | -0.16 | -0.04 | 0.29 | 0.64 |
| Former EL | -0.74 | -0.38 | -0.06 | 0.07 | 0.18 | 0.51 | 0.79 |
| Never EL | -1.19 | -0.48 | -0.12 | 0.00 | 0.13 | 0.50 | 0.94 |
| SWD | -1.00 | -0.85 | -0.43 | -0.30 | -0.15 | 0.27 | 0.83 |
| Not SWD | -1.18 | -0.44 | -0.10 | 0.02 | 0.14 | 0.50 | 0.79 |
| SED | -0.91 | -0.53 | -0.19 | -0.08 | 0.03 | 0.38 | 0.81 |
| Not SED | -0.72 | -0.35 | -0.04 | 0.07 | 0.18 | 0.49 | 0.74 |
| Homeless | -0.99 | -0.69 | -0.27 | -0.11 | 0.02 | 0.40 | 0.70 |
| Not Homeless | -1.19 | -0.49 | -0.13 | -0.01 | 0.11 | 0.46 | 0.94 |
| Foster youth | 0.50 | -0.75 | -0.36 | -0.20 | -0.04 | 0.41 | 0.53 |
| Not foster youth | -1.19 | -0.49 | -0.14 | -0.02 | 0.11 | 0.46 | 0.94 |

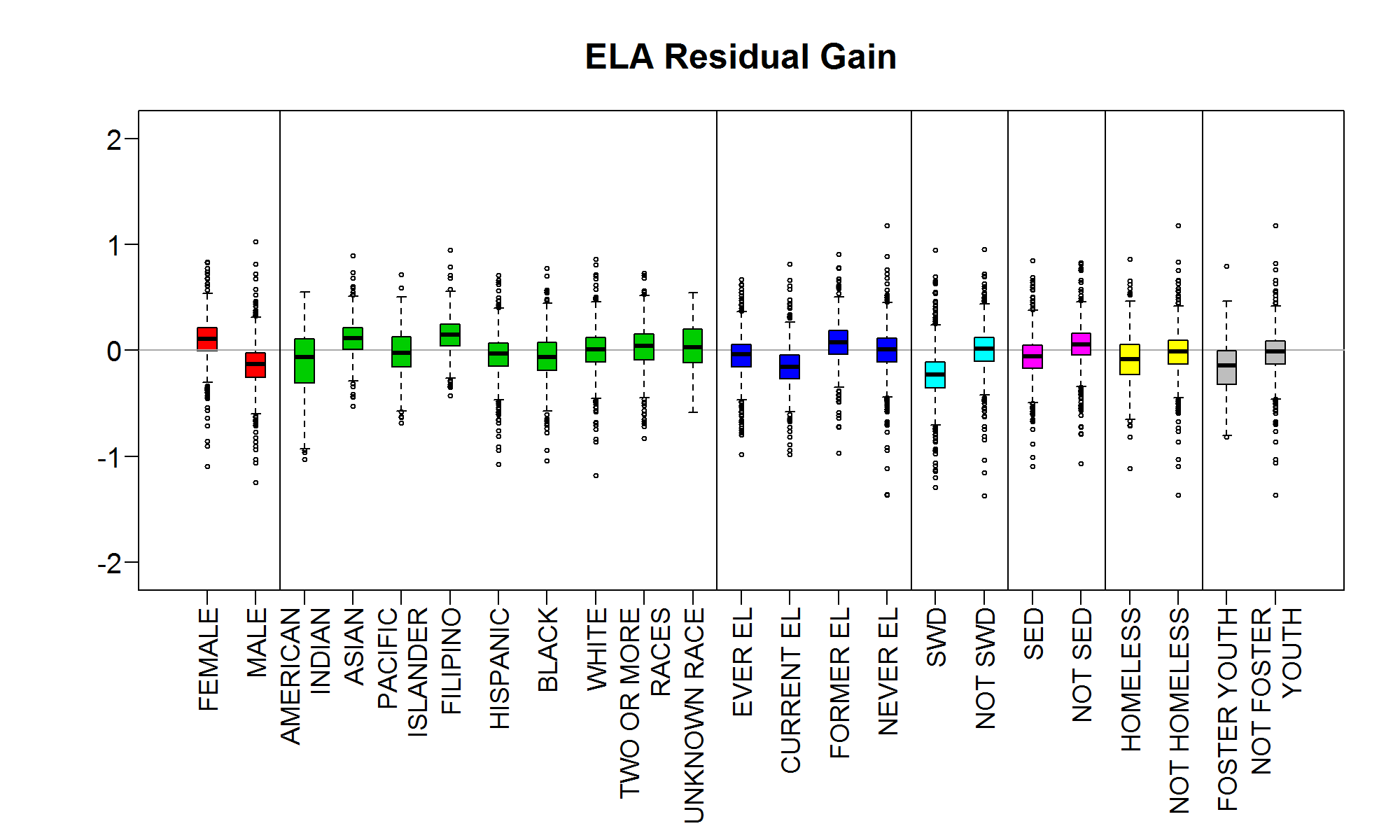


Figure A.4.C. Distributions of student group means for Residual Gain in ELA at the LEA level for the student groups listed along the x-axis. Additionally, student groups within the same group type are organized within vertical lines and color coded (i.e., boxplots for Male and Female are colored red).

Table A.13. Data Supporting Figure A.4.C: Distributions of student group means for Residual Gain in ELA at the LEA level.

| Student Group Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | -1.09 | -0.30 | 0.00 | 0.11 | 0.22 | 0.54 | 0.84 |
| Male | -1.25 | -0.60 | -0.25 | -0.13 | -0.02 | 0.31 | 1.03 |
| American Indian | -1.03 | -0.93 | -0.31 | -0.06 | 0.11 | 0.55 | -0.95 |
| Asian | -0.53 | -0.29 | 0.01 | 0.11 | 0.21 | 0.51 | 0.90 |
| Pacific Islander | -0.68 | -0.57 | -0.15 | -0.02 | 0.13 | 0.51 | 0.72 |
| Filipino | -0.42 | -0.26 | 0.04 | 0.15 | 0.25 | 0.56 | 0.95 |
| Hispanic | -1.07 | -0.47 | -0.15 | -0.03 | 0.07 | 0.40 | 0.71 |
| Black | -1.04 | -0.57 | -0.19 | -0.06 | 0.07 | 0.45 | 0.78 |
| White | -1.18 | -0.46 | -0.11 | 0.01 | 0.12 | 0.46 | 0.87 |
| Two or more races | -0.83 | -0.45 | -0.09 | 0.04 | 0.16 | 0.52 | 0.73 |
| Missing (Unknown) | -0.58 | -0.58 | -0.12 | 0.03 | 0.20 | 0.55 | 0.55 |
| Ever EL | -0.98 | -0.46 | -0.15 | -0.04 | 0.06 | 0.37 | 0.67 |
| Current EL | -0.98 | -0.58 | -0.27 | -0.15 | -0.04 | 0.27 | 0.82 |
| Former EL | -0.97 | -0.35 | -0.04 | 0.07 | 0.19 | 0.51 | 0.91 |
| Never EL | -1.37 | -0.44 | -0.11 | 0.01 | 0.12 | 0.45 | 1.18 |
| SWD | -1.30 | -0.70 | -0.35 | -0.23 | -0.11 | 0.24 | 0.95 |
| Not SWD | -1.38 | -0.42 | -0.10 | 0.02 | 0.12 | 0.44 | 0.96 |
| SED | -1.09 | -0.49 | -0.17 | -0.05 | 0.05 | 0.38 | 0.85 |
| Not SED | -1.07 | -0.34 | -0.04 | 0.06 | 0.16 | 0.46 | 0.83 |
| Homeless | -1.11 | -0.66 | -0.23 | -0.08 | 0.06 | 0.47 | 0.86 |
| Not Homeless | -1.37 | -0.44 | -0.13 | -0.01 | 0.09 | 0.42 | 1.18 |
| Foster youth | -0.82 | -0.80 | -0.32 | -0.14 | 0.00 | 0.47 | 0.80 |
| Not foster youth | -1.37 | -0.46 | -0.13 | -0.01 | 0.09 | 0.42 | 1.18 |

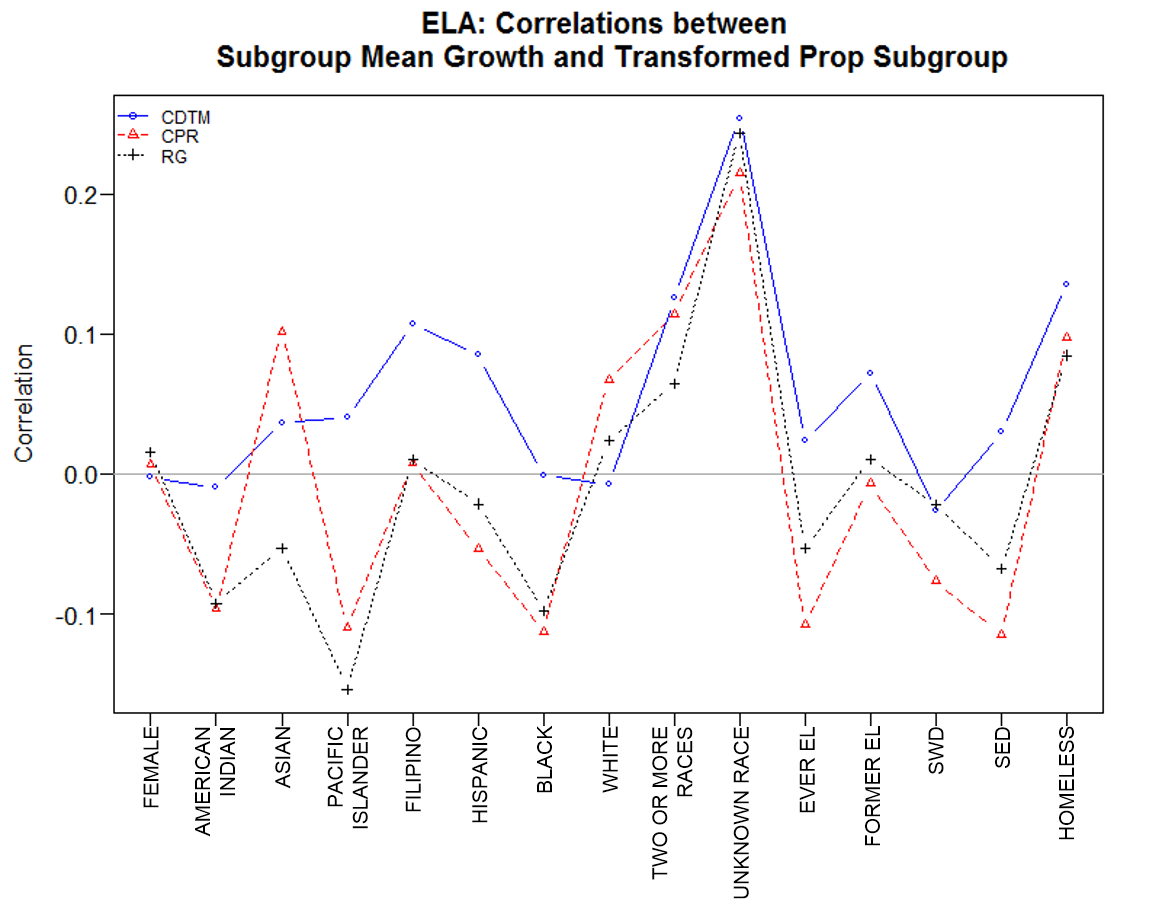


Figure A.5.A. Correlations between each student group mean growth score and (transformed) proportion of students in each student group listed along the x-axis at the school level—ELA. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table A.14. Data Supporting Figure A.5.A: Correlations between each student group mean growth scores and (transformed) proportion of students in each student group at the school level—ELA.

| Student Group Along the X-Axis | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.00 | 0.02 | 0.01 |
| American Indian | -0.01 | -0.09 | -0.10 |
| Asian | 0.04 | -0.05 | 0.10 |
| Pacific Islander | 0.04 | -0.15 | -0.11 |
| Filipino | 0.11 | 0.01 | 0.01 |
| Hispanic | 0.08 | -0.02 | -0.05 |
| Black | 0.00 | -0.10 | -0.11 |
| White | -0.01 | 0.02 | 0.07 |
| Two or more races | 0.13 | 0.06 | 0.11 |
| Missing (Unknown) | 0.25 | 0.24 | 0.22 |
| Ever EL | 0.02 | -0.05 | -0.11 |
| Former EL | 0.07 | 0.01 | -0.01 |
| SWD | -0.03 | -0.02 | -0.08 |
| SED | 0.03 | -0.07 | -0.12 |
| Homeless | 0.14 | 0.08 | 0.10 |

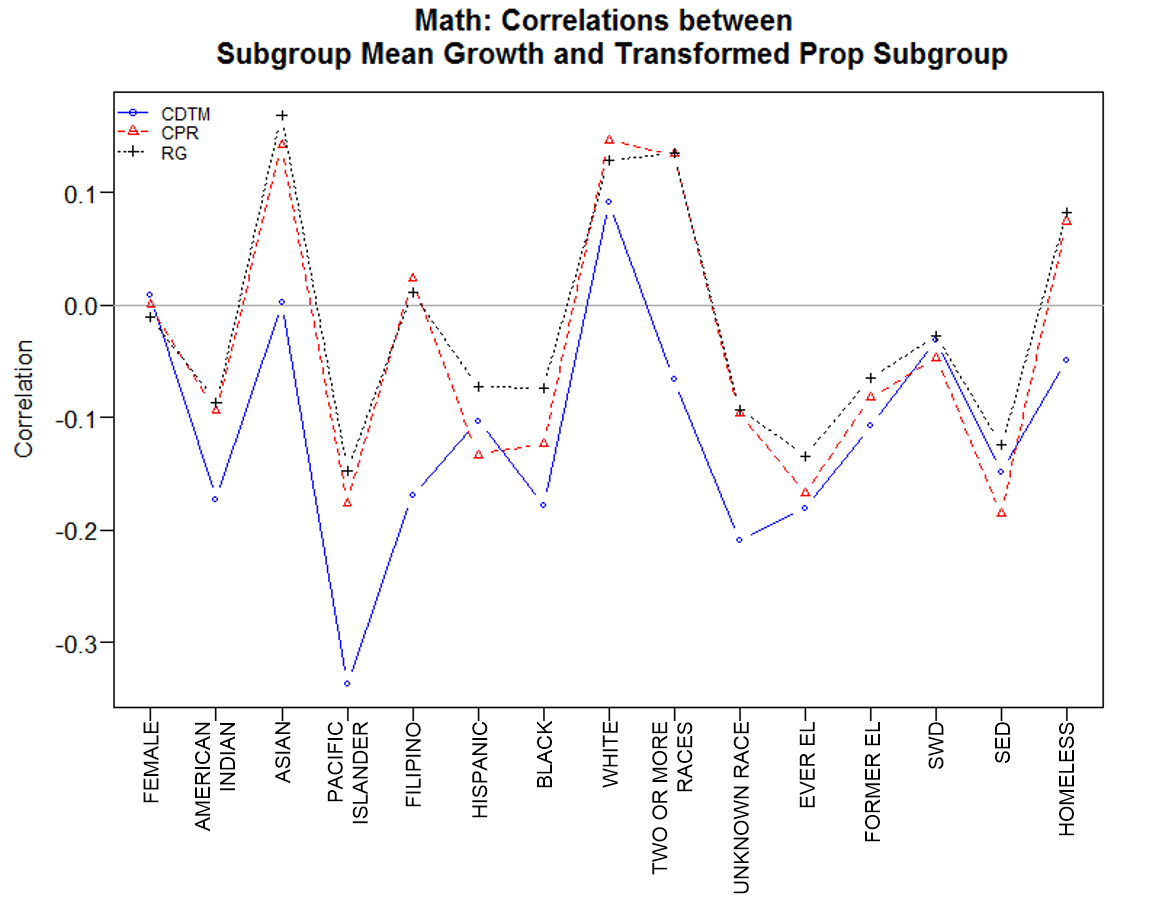


Figure A.5.B. Correlations between each student group mean growth score and (transformed) proportion of students in each student group listed along the x-axis at the school level—mathematics. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table A.15. Data Supporting Figure A.5.B: Correlations between each student group mean growth score and (transformed) proportion of students in each student group at the school level—mathematics.

| Student Group Along the X-Axis | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.01 | -0.01 | 0.00 |
| American Indian | -0.17 | -0.09 | -0.09 |
| Asian | 0.00 | 0.17 | 0.14 |
| Pacific Islander | -0.34 | -0.15 | -0.18 |
| Filipino | -0.17 | 0.01 | 0.02 |
| Hispanic | -0.10 | -0.07 | -0.13 |
| Black | -0.18 | -0.07 | -0.12 |
| White | 0.09 | 0.13 | 0.15 |
| Two or more races | -0.07 | 0.13 | 0.13 |
| Missing (Unknown) | -0.21 | -0.09 | -0.10 |
| Ever EL | -0.18 | -0.13 | -0.17 |
| Former EL | -0.11 | -0.07 | -0.08 |
| SWD | -0.03 | -0.03 | -0.05 |
| SED | -0.15 | -0.12 | -0.19 |
| Homeless | -0.05 | 0.08 | 0.07 |

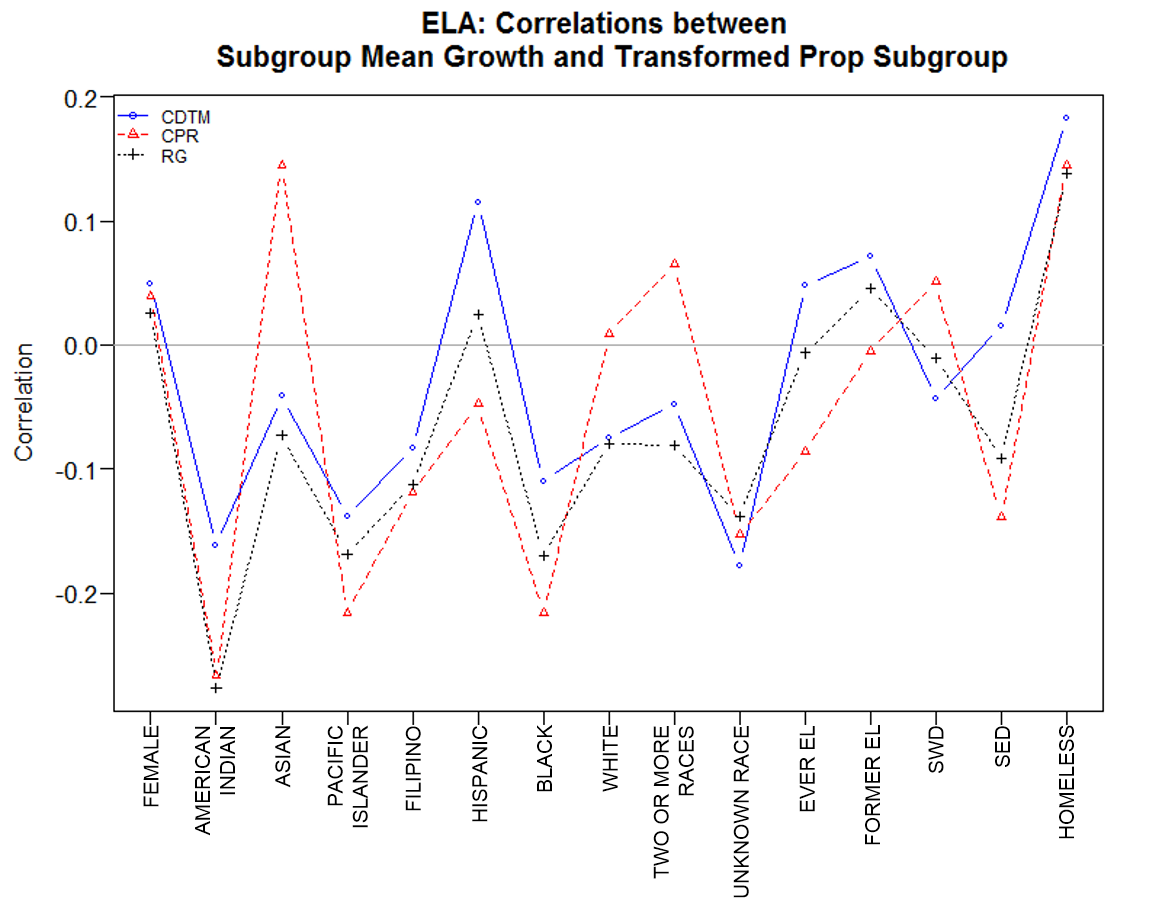


Figure A.6.A. Correlations between each student group mean growth score and (transformed) proportion of students in each student group listed along the x-axis at the LEA level—ELA. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table A.16. Data Supporting Figure A.6.A: Correlations between each student group mean growth score and (transformed) proportion of students in each student group at the LEA level—ELA.

| Student Group Along the X-Axis | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.05 | 0.03 | 0.04 |
| American Indian | -0.16 | -0.28 | -0.27 |
| Asian | -0.04 | -0.07 | 0.14 |
| Pacific Islander | -0.14 | -0.17 | -0.22 |
| Filipino | -0.08 | -0.11 | -0.12 |
| Hispanic | 0.11 | 0.02 | -0.05 |
| Black | -0.11 | -0.17 | -0.22 |
| White | -0.08 | -0.08 | 0.01 |
| Two or more races | -0.05 | -0.08 | 0.06 |
| Missing (Unknown) | -0.18 | -0.14 | -0.15 |
| Ever EL | 0.05 | -0.01 | -0.09 |
| Former EL | 0.07 | 0.05 | 0.00 |
| SWD | -0.04 | -0.01 | 0.05 |
| SED | 0.02 | -0.09 | -0.14 |
| Homeless | 0.18 | 0.14 | 0.15 |

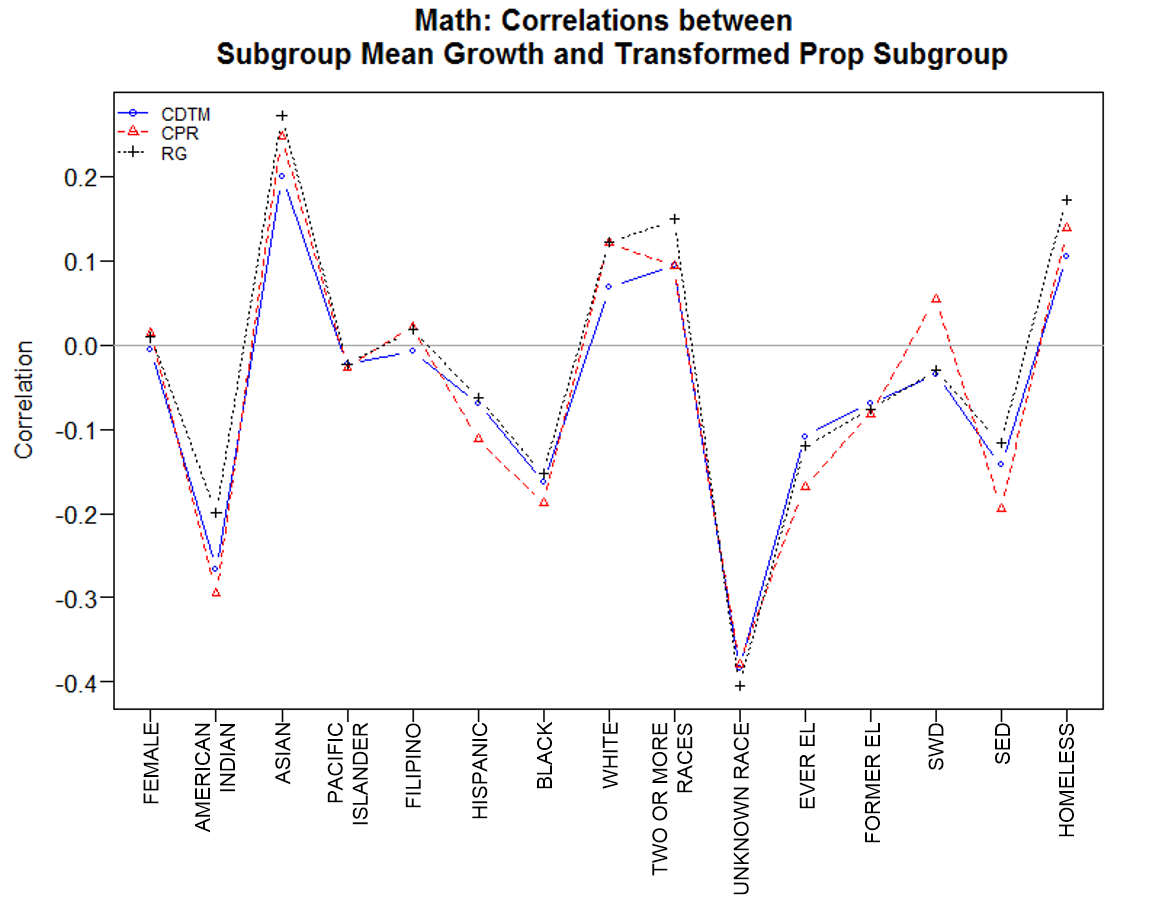


Figure A.6.B. Correlations between each student group mean growth score and (transformed) proportion of students in each student group listed along the x-axis at the LEA level—mathematics. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table A.17. Data Supporting Figure A.6.B: Correlations between each student group mean growth score and (transformed) proportion of students in each student group at the LEA level—mathematics.

| Student Group Along the X-Axis | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | -0.01 | 0.01 | 0.01 |
| American Indian | -0.27 | -0.20 | -0.30 |
| Asian | 0.20 | 0.27 | 0.25 |
| Pacific Islander | -0.02 | -0.02 | -0.03 |
| Filipino | -0.01 | 0.02 | 0.02 |
| Hispanic | -0.07 | -0.06 | -0.11 |
| Black | -0.16 | -0.15 | -0.19 |
| White | 0.07 | 0.12 | 0.12 |
| Two or more races | 0.09 | 0.15 | 0.09 |
| Missing (Unknown) | -0.38 | -0.40 | -0.38 |
| Ever EL | -0.11 | -0.12 | -0.17 |
| Former EL | -0.07 | -0.08 | -0.08 |
| SWD | -0.03 | -0.03 | 0.05 |
| SED | -0.14 | -0.12 | -0.19 |
| Homeless | 0.10 | 0.17 | 0.14 |

# Appendix B: Student Group Comparisons

Table B.1. Data Supporting Figure 1: Correlations between each mean growth statistic at the school level for ELA and each transformed aggregate student demographic variable.

| Student Group | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.02 | 0.07 | 0.03 |
| American Indian | -0.12 | -0.12 | -0.12 |
| Asian | 0.04 | 0.08 | 0.24 |
| Pacific Islander | -0.05 | -0.05 | -0.05 |
| Filipino | 0.02 | 0.08 | 0.11 |
| Hispanic | 0.06 | -0.05 | -0.18 |
| Black | -0.06 | -0.09 | -0.13 |
| White | -0.06 | 0.07 | 0.14 |
| Two or more races | -0.04 | 0.05 | 0.14 |
| Missing (Unknown) | -0.03 | -0.02 | 0.00 |
| Ever EL | 0.09 | -0.07 | -0.13 |
| Former EL | 0.09 | -0.01 | -0.04 |
| SWD | -0.01 | -0.05 | -0.11 |
| SED | -0.03 | -0.19 | -0.32 |
| Homeless | 0.00 | -0.05 | -0.11 |
| Foster youth | -0.05 | -0.08 | -0.13 |

Table B.2. Data Supporting Figure 2: Correlations between each mean growth statistic at the school level for mathematics and each transformed aggregate student demographic variable.

| Student Group | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.03 | 0.00 | 0.03 |
| American Indian | -0.03 | -0.06 | -0.07 |
| Asian | 0.19 | 0.24 | 0.25 |
| Pacific Islander | -0.01 | -0.02 | -0.03 |
| Filipino | 0.09 | 0.07 | 0.10 |
| Hispanic | -0.19 | -0.20 | -0.25 |
| Black | -0.14 | -0.15 | -0.16 |
| White | 0.18 | 0.15 | 0.21 |
| Two or more races | 0.10 | 0.12 | 0.17 |
| Missing (Unknown) | 0.01 | 0.01 | 0.01 |
| Ever EL | -0.18 | -0.10 | -0.17 |
| Former EL | -0.07 | -0.04 | -0.09 |
| SWD | -0.08 | -0.05 | -0.07 |
| SED | -0.29 | -0.28 | -0.36 |
| Homeless | -0.10 | -0.09 | -0.12 |
| Foster youth | -0.08 | -0.11 | -0.13 |

Table B.3. Data Supporting Figure 3: Correlations between each mean growth statistic at the LEA level for ELA and each transformed aggregate student demographic variable.

| Student Group | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.02 | 0.08 | 0.05 |
| American Indian | -0.23 | -0.29 | -0.28 |
| Asian | 0.08 | 0.17 | 0.28 |
| Pacific Islander | 0.03 | 0.05 | 0.06 |
| Filipino | 0.03 | 0.10 | 0.11 |
| Hispanic | 0.10 | 0.02 | -0.12 |
| Black | -0.05 | -0.02 | -0.06 |
| White | -0.08 | 0.00 | 0.09 |
| Two or more races | -0.03 | 0.04 | 0.15 |
| Missing (Unknown) | 0.04 | 0.07 | 0.08 |
| Ever EL | 0.08 | 0.01 | -0.08 |
| Former EL | 0.12 | 0.10 | 0.04 |
| SWD | -0.06 | -0.02 | -0.06 |
| SED | -0.03 | -0.20 | -0.36 |
| Homeless | 0.02 | -0.02 | -0.08 |
| Foster youth | 0.02 | -0.02 | -0.07 |

Table B.4. Data Supporting Figure 4: Correlations between each mean growth statistic at the LEA level for mathematics and each transformed aggregate student demographic variable.

| Student Group | CDTM | RG | CPR |
| --- | --- | --- | --- |
| Female | 0.04 | 0.01 | 0.06 |
| American Indian | -0.23 | -0.21 | -0.24 |
| Asian | 0.24 | 0.22 | 0.27 |
| Pacific Islander | 0.02 | 0.00 | 0.01 |
| Filipino | 0.08 | 0.04 | 0.08 |
| Hispanic | -0.17 | -0.20 | -0.23 |
| Black | -0.10 | -0.13 | -0.11 |
| White | 0.15 | 0.18 | 0.21 |
| Two or more races | 0.17 | 0.18 | 0.20 |
| Missing (Unknown) | 0.01 | 0.01 | 0.03 |
| Ever EL | -0.13 | -0.14 | -0.17 |
| Former EL | -0.02 | -0.06 | -0.05 |
| SWD | -0.07 | -0.09 | -0.05 |
| SED | -0.34 | -0.32 | -0.40 |
| Homeless | -0.09 | -0.10 | -0.13 |
| Foster youth | -0.13 | -0.14 | -0.16 |

Table B.5. Data Supporting Top Panel of Figure 5: Percentiles of Met Standard thresholds for each grade level and content area.

| Grade | ELA | Math |
| --- | --- | --- |
| 4 | 55.15 | 60.76 |
| 5 | 50.19 | 66.04 |
| 6 | 51.66 | 63.64 |
| 7 | 50.78 | 62.36 |
| 8 | 49.45 | 62.61 |

Table B.6. Data Supporting Bottom Panel of Figure 5: Differences in percentiles of thresholds for each pair of adjacent grade levels.

| Grade | ELA | Math |
| --- | --- | --- |
| 5-4 | -4.96 | 5.28 |
| 6-5 | 1.47 | -2.40 |
| 7-6 | -0.87 | -1.28 |
| 8-7 | -1.33 | 0.25 |

# Appendix C: Student Group Comparisons

Table C.1. Numbers of schools and LEAs with at least 10 students in *both* student groups within a contrast\*

| Category | Contrast | N Schools | N LEAs |
| --- | --- | --- | --- |
| Gender | Female–Male | 7,168 | 799 |
| Race/Ethnicity | White–American Indian | 122 | 251 |
| Race/Ethnicity | White–Asian | 2,504 | 404 |
| Race/Ethnicity | White–Pacific Islander | 120 | 192 |
| Race/Ethnicity | White–Filipino | 1,097 | 313 |
| Race/Ethnicity | White–Hispanic | 4,820 | 693 |
| Race/Ethnicity | White–Black | 1,880 | 361 |
| Race/Ethnicity | White–Two or more races | 2,009 | 444 |
| Race/Ethnicity | White–Missing (Unknown) | 72 | 70 |
| English Learner status | Never EL–Current EL | 5,676 | 624 |
| English Learner status | Never EL–Former EL | 5,699 | 594 |
| English Learner status | Never EL–Ever EL | 6,447 | 673 |
| Disability status | SWD–Not SWD | 6,478 | 690 |
| Economically disadvantaged status | SED–Not SED | 5,955 | 729 |
| Homeless status | Homeless–Not homeless | 1,635 | 352 |
| Foster care status | Foster Youth–Not Foster Youth | 2 | 112 |

\*Note: These N counts of schools and LEAs are used for each subsequent graph in this appendix.

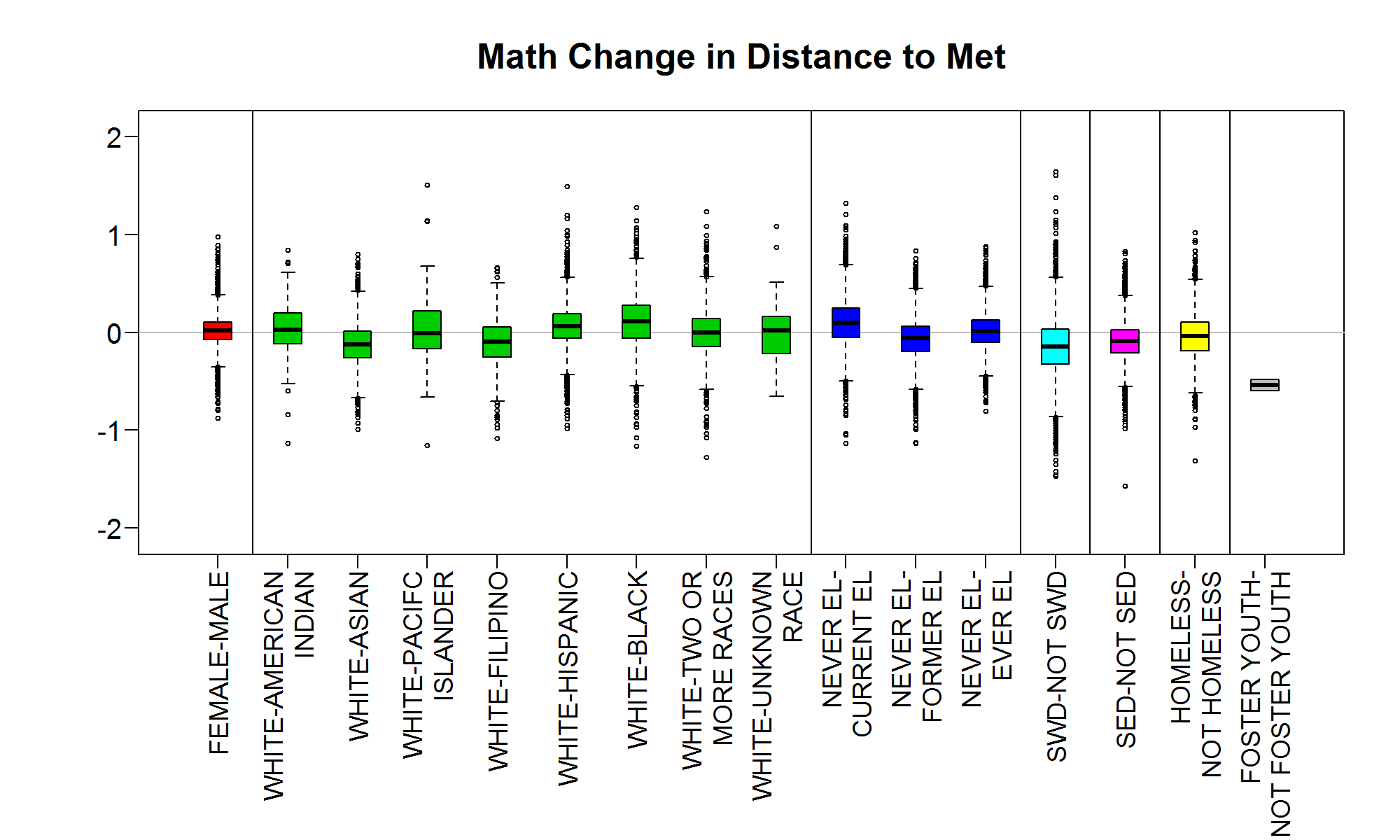


Figure C.1.A. Distributions of group differences (or gaps) for Change in Distance to Met in mathematics at the school for the student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.2. Data Supporting Figure C.1.A: Distributions of group differences (or gaps) for Change in Distance to Met in mathematics at the school level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.88 | -0.35 | -0.08 | 0.02 | 0.11 | 0.38 | 0.98 |
| White–American Indian | -1.13 | -0.52 | -0.11 | 0.03 | 0.20 | 0.61 | 0.85 |
| White–Asian | -0.99 | -0.67 | -0.26 | -0.12 | 0.01 | 0.42 | 0.80 |
| White–Pacific Islander | -1.16 | -0.66 | -0.16 | -0.01 | 0.22 | 0.67 | 1.51 |
| White–Filipino | -1.08 | -0.70 | -0.25 | -0.10 | 0.05 | 0.50 | 0.67 |
| White–Hispanic | -0.98 | -0.43 | -0.06 | 0.07 | 0.19 | 0.56 | 1.49 |
| White–Black | -1.16 | -0.54 | -0.06 | 0.11 | 0.27 | 0.76 | 1.28 |
| White–Two or more races | -1.28 | -0.58 | -0.15 | 0.00 | 0.14 | 0.57 | 1.24 |
| White–Missing (Unknown) | 0.87 | -0.65 | -0.21 | 0.02 | 0.16 | 0.51 | 1.09 |
| Never EL–Current EL | -1.13 | -0.49 | -0.05 | 0.10 | 0.25 | 0.69 | 1.33 |
| Never EL–Former EL | -1.13 | -0.58 | -0.19 | -0.06 | 0.07 | 0.45 | 0.84 |
| Never EL–Ever EL | -0.80 | -0.44 | -0.10 | 0.01 | 0.13 | 0.47 | 0.88 |
| SWD–Not SWD | -3.21 | -0.86 | -0.33 | -0.15 | 0.03 | 0.56 | 1.64 |
| SED–Not SED | -1.57 | -0.56 | -0.21 | -0.09 | 0.03 | 0.37 | 0.83 |
| Homeless–Not homeless | -1.31 | -0.62 | -0.19 | -0.04 | 0.11 | 0.55 | 1.02 |
| Foster youth–Not foster youth | -0.60 | -0.60 | -0.60 | -0.54 | -0.48 | -0.48 | -0.48 |

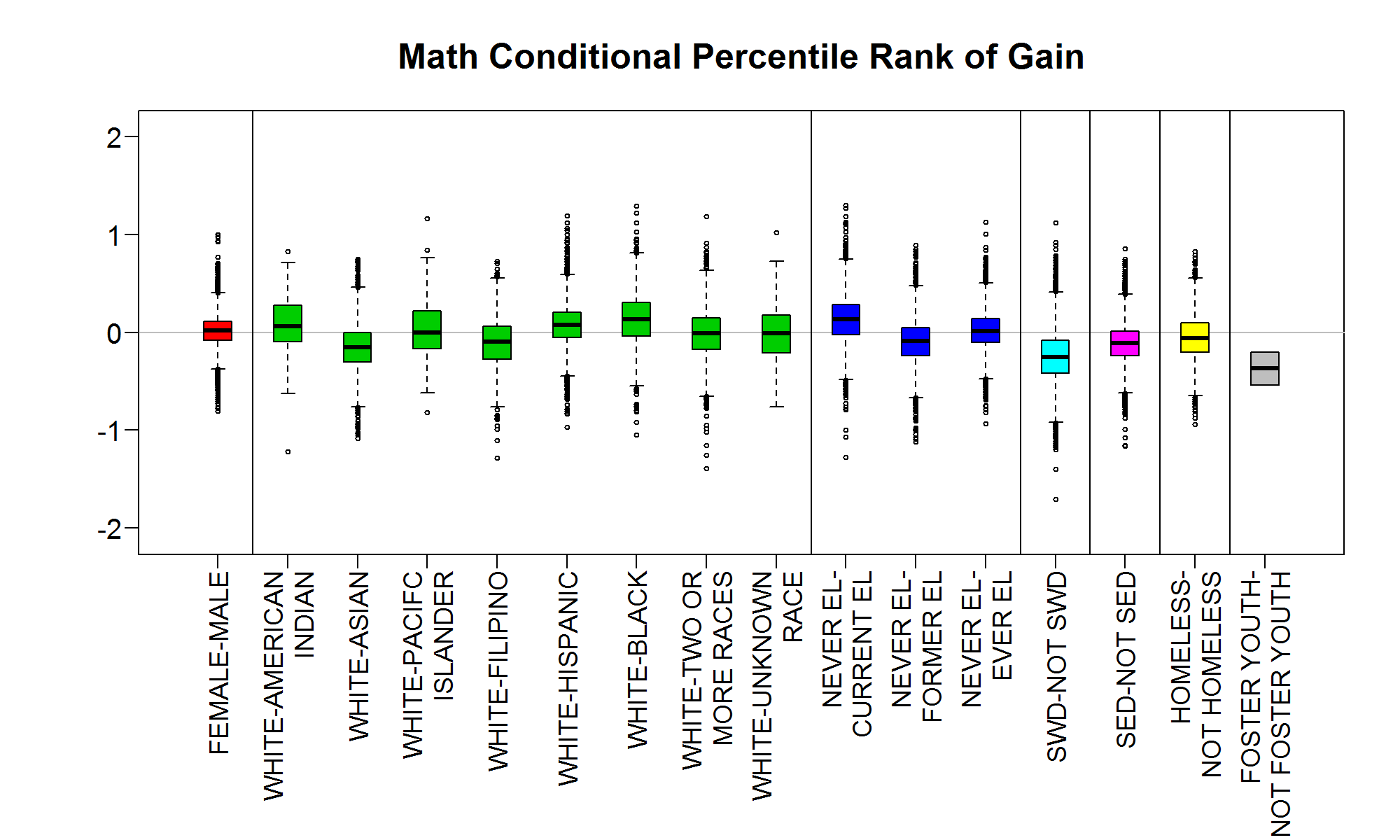


Figure C.1.B. Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in mathematics at the school level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.3. Data Supporting Figure C.1.B: Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in mathematics at the school level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.81 | -0.37 | -0.08 | 0.02 | 0.11 | 0.40 | 1.00 |
| White–American Indian | -1.22 | -0.62 | -0.09 | 0.06 | 0.27 | 0.71 | 0.83 |
| White–Asian | -1.08 | -0.76 | -0.30 | -0.15 | 0.00 | 0.46 | 0.75 |
| White–Pacific Islander | -0.81 | -0.62 | -0.17 | 0.00 | 0.22 | 0.77 | 1.16 |
| White–Filipino | -1.28 | -0.76 | -0.27 | -0.10 | 0.06 | 0.55 | 0.73 |
| White–Hispanic | -0.97 | -0.44 | -0.05 | 0.08 | 0.21 | 0.59 | 1.19 |
| White–Black | -1.04 | -0.55 | -0.04 | 0.13 | 0.30 | 0.81 | 1.29 |
| White–Two or more races | -1.39 | -0.65 | -0.17 | -0.01 | 0.15 | 0.63 | 1.19 |
| White–Missing (Unknown) | 1.02 | -0.76 | -0.21 | -0.01 | 0.18 | 0.73 | 1.02 |
| Never EL–Current EL | -1.27 | -0.48 | -0.02 | 0.13 | 0.29 | 0.75 | 1.30 |
| Never EL–Former EL | -1.12 | -0.67 | -0.24 | -0.08 | 0.05 | 0.48 | 0.89 |
| Never EL–Ever EL | -0.93 | -0.47 | -0.11 | 0.01 | 0.14 | 0.51 | 1.13 |
| SWD–Not SWD | -1.71 | -0.92 | -0.42 | -0.25 | -0.08 | 0.41 | 1.12 |
| SED–Not SED | -1.16 | -0.62 | -0.24 | -0.11 | 0.01 | 0.39 | 0.85 |
| Homeless–Not homeless | -0.94 | -0.65 | -0.20 | -0.06 | 0.10 | 0.56 | 0.83 |
| Foster youth–Not foster youth | -0.54 | -0.54 | -0.54 | -0.37 | -0.20 | -0.20 | -0.20 |

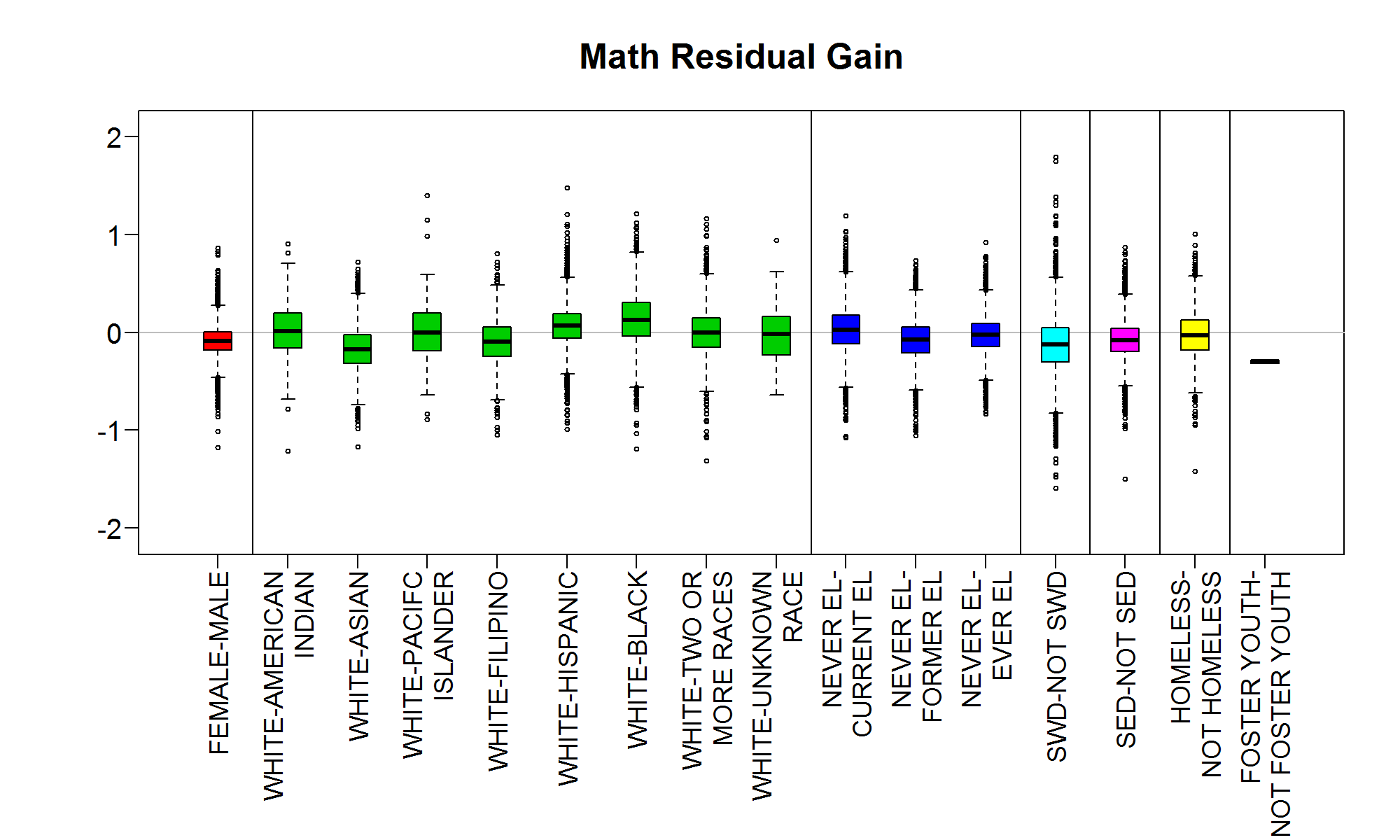


Figure C.1.C. Distributions of group differences (or gaps) for Residual Gain in mathematics at the school level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.4. Data Supporting Figure C.1.C: Distributions of group differences (or gaps) for Residual Gain in mathematics at the school level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -1.18 | -0.46 | -0.18 | -0.09 | 0.00 | 0.28 | 0.87 |
| White–American Indian | -1.21 | -0.68 | -0.16 | 0.01 | 0.20 | 0.71 | 0.90 |
| White–Asian | -1.17 | -0.74 | -0.31 | -0.17 | -0.03 | 0.40 | 0.72 |
| White–Pacific Islander | -0.89 | -0.64 | -0.19 | 0.00 | 0.20 | 0.59 | 1.40 |
| White–Filipino | -1.04 | -0.69 | -0.24 | -0.09 | 0.06 | 0.49 | 0.81 |
| White–Hispanic | -0.99 | -0.43 | -0.06 | 0.07 | 0.19 | 0.56 | 1.48 |
| White–Black | -1.19 | -0.56 | -0.04 | 0.13 | 0.31 | 0.82 | 1.22 |
| White–Two or more races | -1.31 | -0.60 | -0.15 | 0.00 | 0.15 | 0.60 | 1.17 |
| White–Missing (Unknown) | 0.95 | -0.64 | -0.23 | -0.02 | 0.17 | 0.62 | 0.95 |
| Never EL–Current EL | -1.07 | -0.56 | -0.12 | 0.03 | 0.18 | 0.62 | 1.20 |
| Never EL–Former EL | -1.05 | -0.59 | -0.21 | -0.07 | 0.05 | 0.44 | 0.73 |
| Never EL–Ever EL | -0.83 | -0.49 | -0.14 | -0.03 | 0.09 | 0.43 | 0.92 |
| SWD–Not SWD | -3.25 | -0.83 | -0.30 | -0.12 | 0.05 | 0.57 | 1.80 |
| SED–Not SED | -1.49 | -0.55 | -0.20 | -0.08 | 0.04 | 0.39 | 0.87 |
| Homeless–Not homeless | -1.42 | -0.62 | -0.18 | -0.03 | 0.13 | 0.58 | 1.00 |
| Foster youth–Not foster youth | -0.32 | -0.32 | -0.32 | -0.30 | -0.28 | -0.28 | -0.28 |

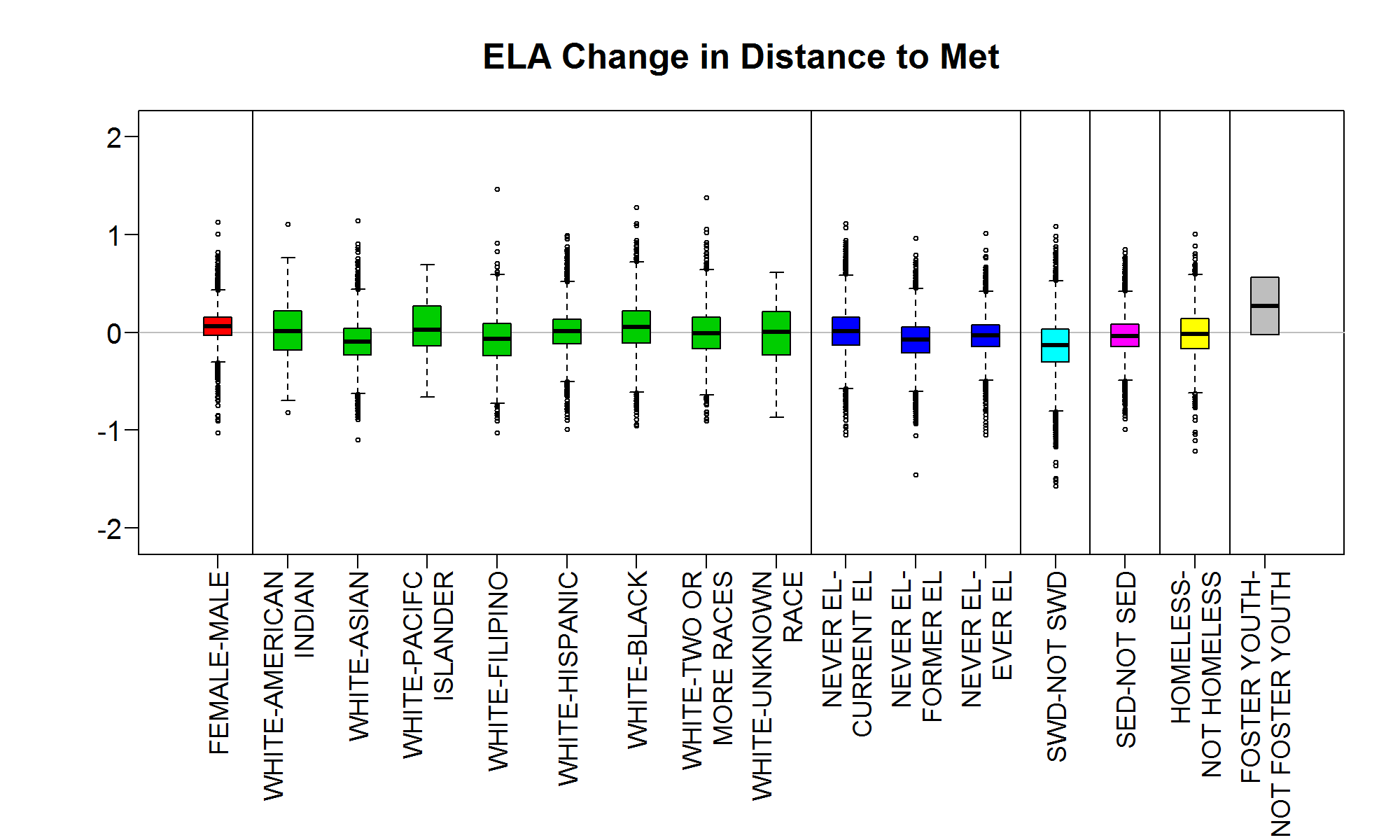


Figure C.2.A. Distributions of group differences (or gaps) for Change in Distance to Met in ELA at the school level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.5. Data Supporting Figure C.2.A: Distributions of group differences (or gaps) for Change in Distance to Met in ELA at the school level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -1.03 | -0.31 | -0.03 | 0.06 | 0.16 | 0.43 | 1.13 |
| White–American Indian | -0.82 | -0.70 | -0.18 | 0.01 | 0.22 | 0.76 | 1.11 |
| White–Asian | -1.10 | -0.63 | -0.23 | -0.10 | 0.04 | 0.44 | 1.14 |
| White–Pacific Islander | -0.66 | -0.66 | -0.14 | 0.03 | 0.27 | 0.70 | 0.70 |
| White–Filipino | -1.03 | -0.73 | -0.24 | -0.07 | 0.09 | 0.59 | 1.47 |
| White–Hispanic | -0.99 | -0.50 | -0.12 | 0.01 | 0.14 | 0.52 | 0.99 |
| White–Black | -0.95 | -0.61 | -0.11 | 0.05 | 0.22 | 0.72 | 1.28 |
| White–Two or more races | -0.90 | -0.64 | -0.17 | -0.01 | 0.16 | 0.64 | 1.38 |
| White–Missing (Unknown) | -0.87 | -0.87 | -0.23 | 0.01 | 0.21 | 0.61 | 0.61 |
| Never EL–Current EL | -1.05 | -0.57 | -0.13 | 0.01 | 0.16 | 0.59 | 1.12 |
| Never EL–Former EL | -1.45 | -0.61 | -0.21 | -0.07 | 0.06 | 0.45 | 0.97 |
| Never EL–Ever EL | -1.04 | -0.49 | -0.15 | -0.03 | 0.08 | 0.42 | 1.02 |
| SWD–Not SWD | -3.46 | -0.80 | -0.31 | -0.13 | 0.03 | 0.53 | 1.09 |
| SED–Not SED | -0.99 | -0.49 | -0.15 | -0.04 | 0.08 | 0.42 | 0.85 |
| Homeless–Not homeless | -1.21 | -0.61 | -0.16 | -0.02 | 0.14 | 0.59 | 1.01 |
| Foster youth–Not foster youth | -0.02 | -0.02 | -0.02 | 0.27 | 0.56 | 0.56 | 0.56 |

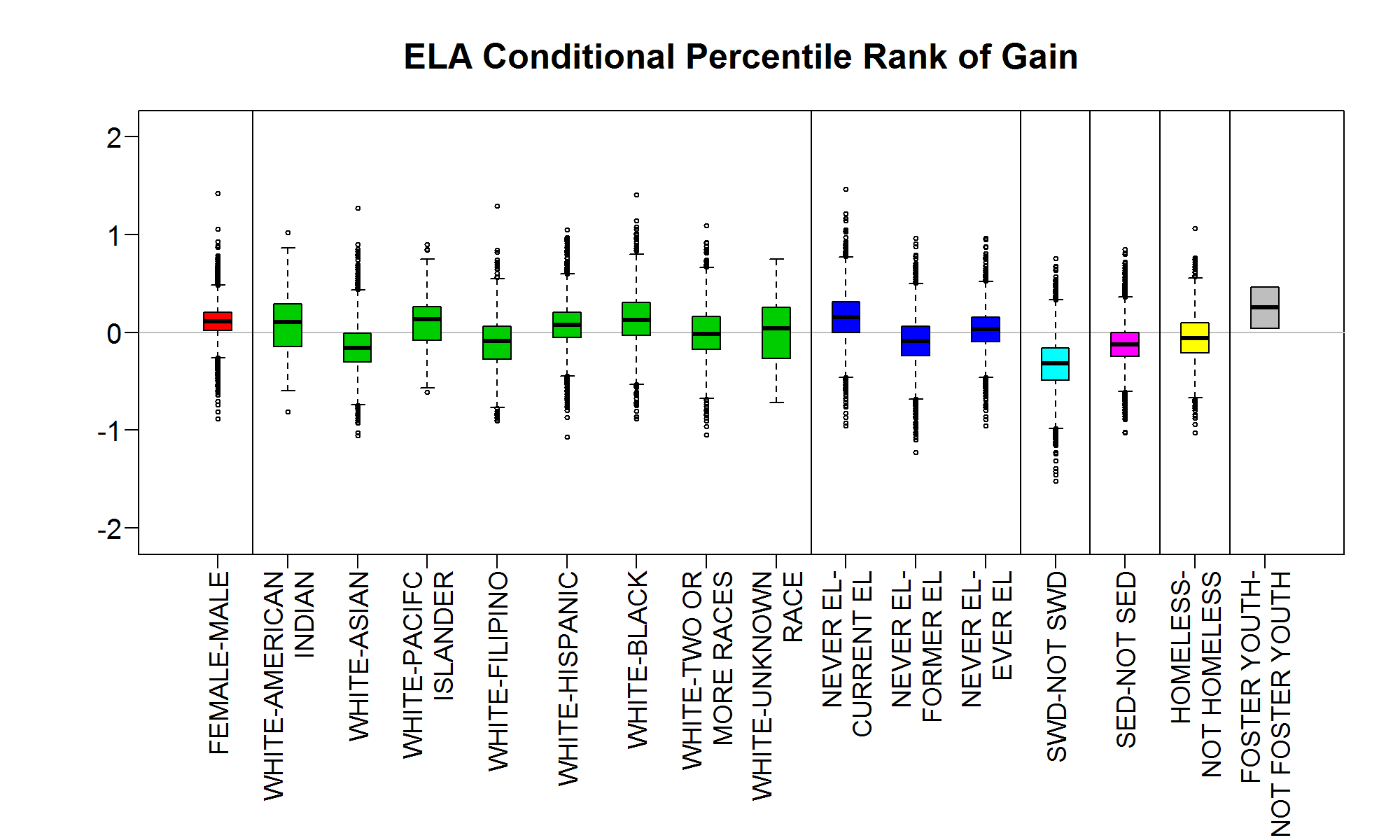


Figure C.2.B. Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in ELA at the school level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.6. Data Supporting Figure C.2.B: Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in ELA at the school level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.88 | -0.26 | 0.02 | 0.11 | 0.21 | 0.48 | 1.42 |
| White–American Indian | -0.81 | -0.59 | -0.14 | 0.10 | 0.29 | 0.86 | 1.02 |
| White–Asian | -1.06 | -0.74 | -0.30 | -0.16 | -0.01 | 0.43 | 1.27 |
| White–Pacific Islander | -0.61 | -0.57 | -0.08 | 0.13 | 0.27 | 0.75 | 0.90 |
| White–Filipino | -0.90 | -0.77 | -0.27 | -0.09 | 0.06 | 0.55 | 1.30 |
| White–Hispanic | -1.07 | -0.44 | -0.05 | 0.08 | 0.21 | 0.60 | 1.05 |
| White–Black | -0.88 | -0.53 | -0.03 | 0.13 | 0.30 | 0.80 | 1.41 |
| White–Two or more races | -1.05 | -0.67 | -0.17 | -0.02 | 0.16 | 0.66 | 1.09 |
| White–Missing (Unknown) | -0.72 | -0.72 | -0.27 | 0.04 | 0.25 | 0.75 | 0.75 |
| Never EL–Current EL | -0.96 | -0.46 | 0.00 | 0.16 | 0.31 | 0.77 | 1.47 |
| Never EL–Former EL | -1.22 | -0.68 | -0.24 | -0.08 | 0.06 | 0.50 | 0.97 |
| Never EL–Ever EL | -0.96 | -0.46 | -0.09 | 0.03 | 0.15 | 0.52 | 0.96 |
| SWD–Not SWD | -1.52 | -0.98 | -0.49 | -0.32 | -0.16 | 0.33 | 0.76 |
| SED–Not SED | -1.03 | -0.61 | -0.24 | -0.12 | 0.00 | 0.36 | 0.85 |
| Homeless–Not homeless | -1.03 | -0.66 | -0.21 | -0.06 | 0.10 | 0.56 | 1.06 |
| Foster youth–Not foster youth | 0.04 | 0.04 | 0.04 | 0.25 | 0.46 | 0.46 | 0.46 |

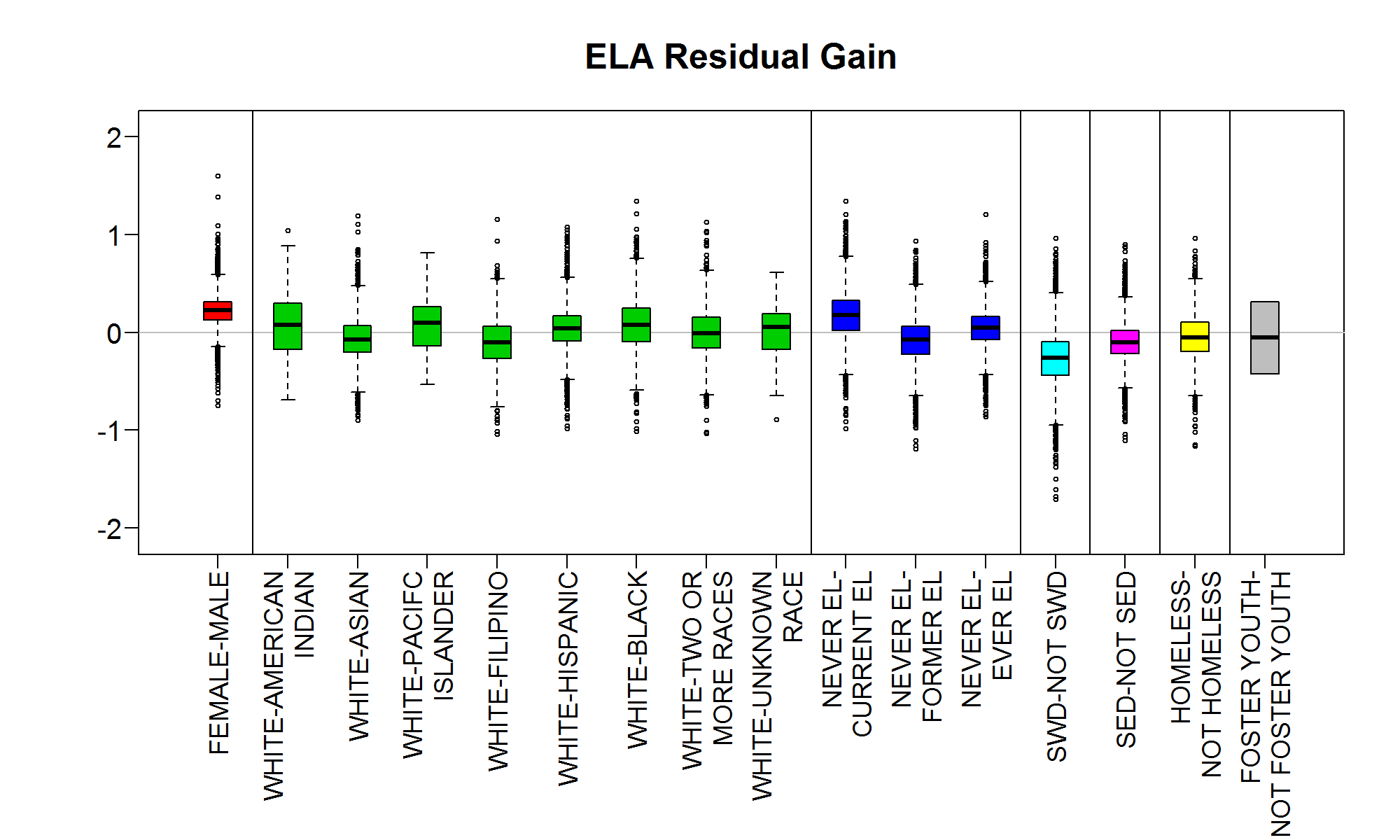


Figure C.2.C. Distributions of group differences (or gaps) for Residual Gain in ELA at the school level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.7. Data Supporting Figure C.2.C: Distributions of group differences (or gaps) for Residual Gain in ELA at the school level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.75 | -0.15 | 0.13 | 0.22 | 0.32 | 0.59 | 1.60 |
| White–American Indian | 1.04 | -0.69 | -0.17 | 0.07 | 0.30 | 0.89 | 1.04 |
| White–Asian | -0.90 | -0.61 | -0.20 | -0.07 | 0.07 | 0.48 | 1.19 |
| White–Pacific Islander | -0.53 | -0.53 | -0.14 | 0.10 | 0.26 | 0.81 | 0.81 |
| White–Filipino | -1.04 | -0.76 | -0.27 | -0.10 | 0.06 | 0.55 | 1.15 |
| White–Hispanic | -0.98 | -0.48 | -0.09 | 0.04 | 0.17 | 0.56 | 1.08 |
| White–Black | -1.01 | -0.59 | -0.09 | 0.08 | 0.25 | 0.75 | 1.35 |
| White–Two or more races | -1.03 | -0.64 | -0.16 | -0.01 | 0.16 | 0.63 | 1.13 |
| White–Missing (Unknown) | -0.89 | -0.65 | -0.17 | 0.06 | 0.19 | 0.62 | -0.89 |
| Never EL–Current EL | -0.98 | -0.43 | 0.02 | 0.18 | 0.33 | 0.78 | 1.34 |
| Never EL–Former EL | -1.19 | -0.65 | -0.22 | -0.07 | 0.06 | 0.49 | 0.94 |
| Never EL–Ever EL | -0.86 | -0.43 | -0.08 | 0.05 | 0.16 | 0.52 | 1.21 |
| SWD–Not SWD | -3.93 | -0.95 | -0.44 | -0.26 | -0.10 | 0.41 | 0.97 |
| SED–Not SED | -1.10 | -0.57 | -0.22 | -0.10 | 0.02 | 0.37 | 0.90 |
| Homeless–Not homeless | -1.16 | -0.65 | -0.20 | -0.05 | 0.10 | 0.55 | 0.97 |
| Foster youth–Not foster youth | -0.42 | -0.42 | -0.42 | -0.06 | 0.31 | 0.31 | 0.31 |

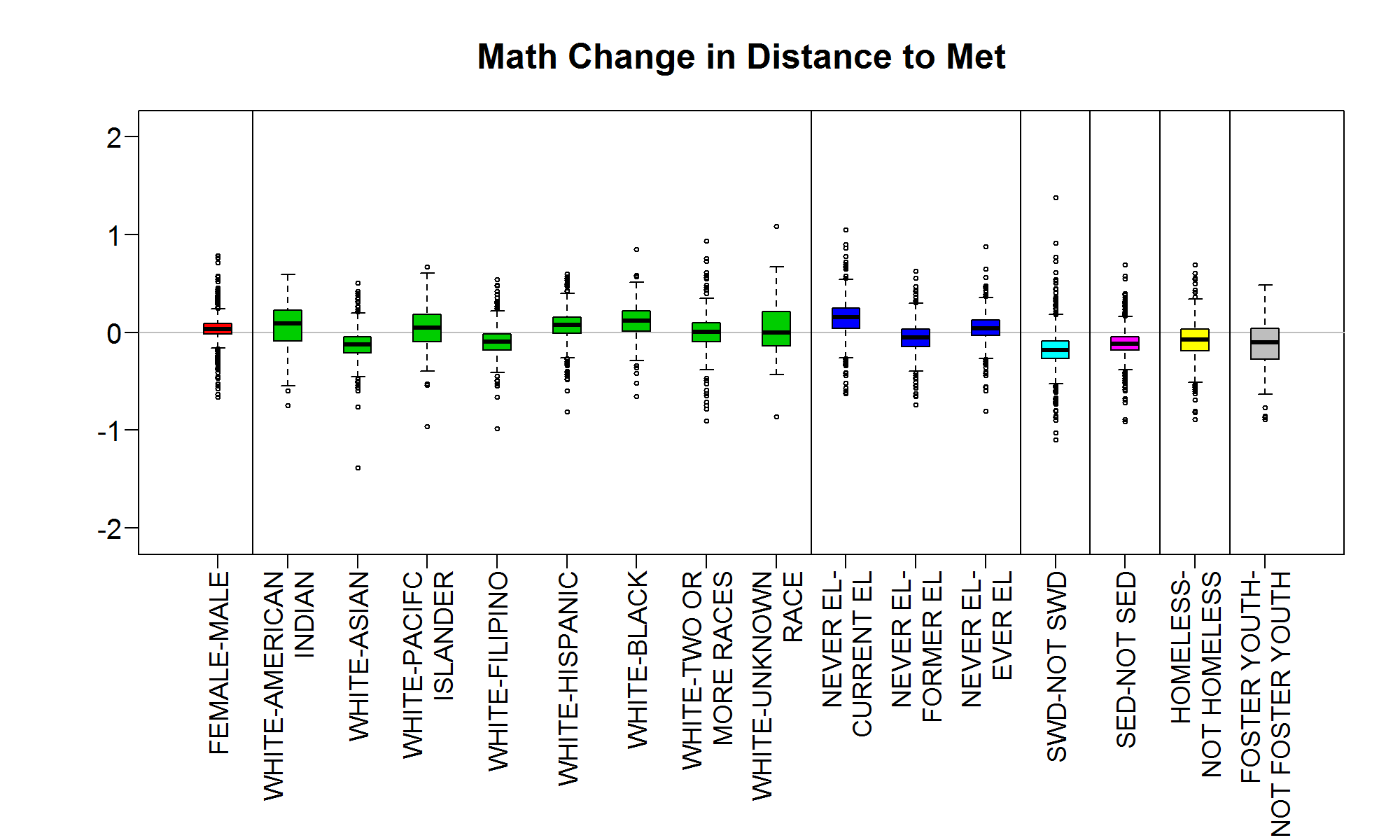


Figure C.3.A. Distributions of group differences (or gaps) for Change in Distance to Met in mathematics at the LEA level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.8. Data Supporting Figure C.3.A: Distributions of group differences (or gaps) for Change in Distance to Met in mathematics at the LEA level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.66 | -0.16 | -0.01 | 0.03 | 0.09 | 0.24 | 0.79 |
| White–American Indian | -0.75 | -0.55 | -0.09 | 0.09 | 0.22 | 0.59 | -0.60 |
| White–Asian | -1.39 | -0.45 | -0.21 | -0.12 | -0.04 | 0.20 | 0.50 |
| White–Pacific Islander | -0.96 | -0.39 | -0.10 | 0.05 | 0.18 | 0.61 | 0.67 |
| White–Filipino | -0.98 | -0.41 | -0.18 | -0.10 | -0.01 | 0.22 | 0.54 |
| White–Hispanic | -0.81 | -0.26 | -0.01 | 0.08 | 0.16 | 0.40 | 0.60 |
| White–Black | -0.65 | -0.29 | 0.01 | 0.12 | 0.22 | 0.52 | 0.85 |
| White–Two or more races | -0.90 | -0.38 | -0.09 | 0.01 | 0.10 | 0.35 | 0.93 |
| White–Missing (Unknown) | -0.86 | -0.43 | -0.14 | 0.00 | 0.22 | 0.67 | 1.09 |
| Never EL–Current EL | -0.62 | -0.26 | 0.04 | 0.15 | 0.25 | 0.54 | 1.05 |
| Never EL–Former EL | -0.74 | -0.40 | -0.14 | -0.05 | 0.04 | 0.30 | 0.63 |
| Never EL–Ever EL | -0.80 | -0.27 | -0.03 | 0.04 | 0.13 | 0.36 | 0.88 |
| SWD–Not SWD | -1.10 | -0.53 | -0.26 | -0.18 | -0.08 | 0.18 | 1.38 |
| SED–Not SED | -0.91 | -0.38 | -0.18 | -0.11 | -0.04 | 0.17 | 0.69 |
| Homeless–Not homeless | -0.89 | -0.51 | -0.19 | -0.07 | 0.03 | 0.34 | 0.69 |
| Foster youth–Not foster youth | -0.89 | -0.63 | -0.28 | -0.10 | 0.04 | 0.49 | -0.77 |

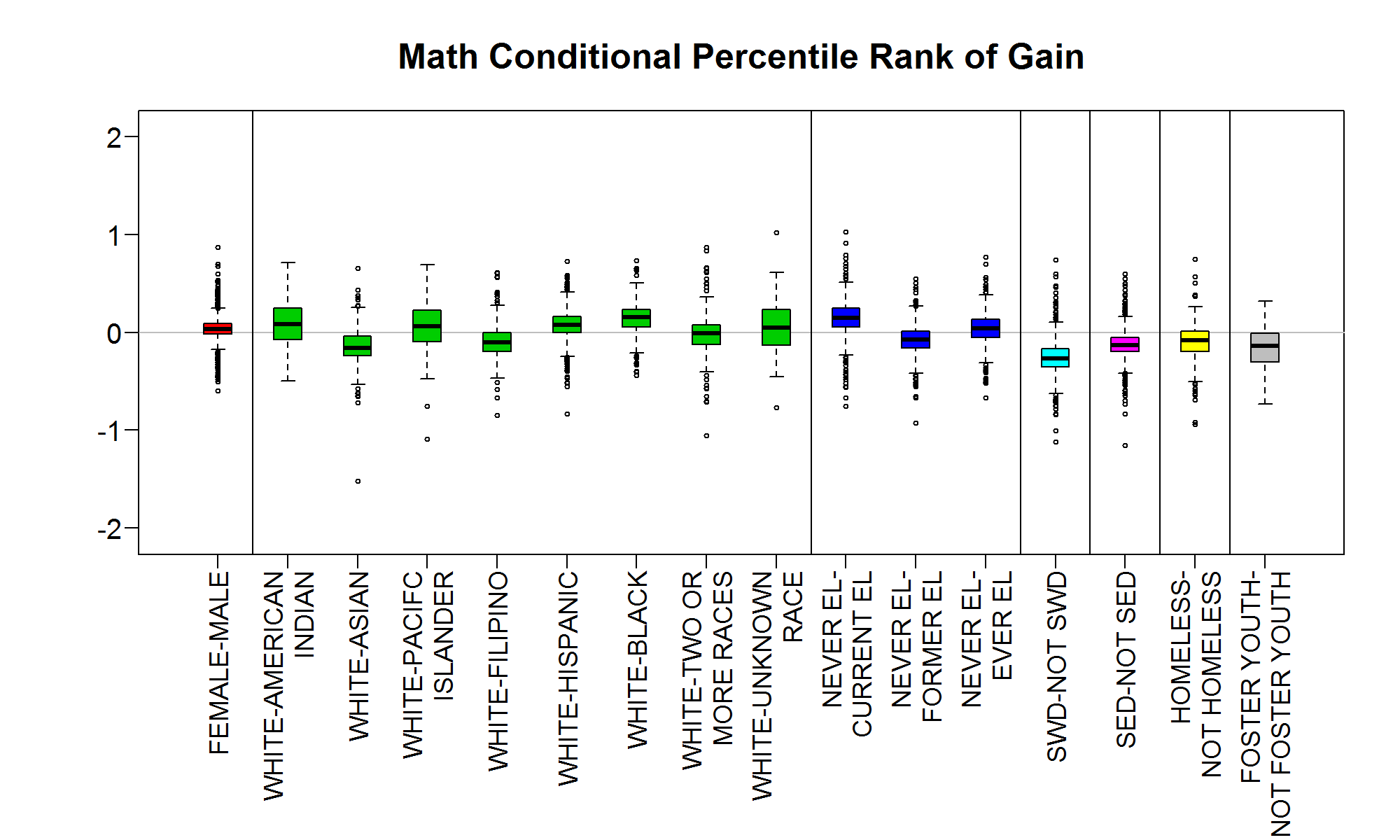


Figure C.3.B. Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in mathematics at the LEA level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.9. Data Supporting Figure C.3.B: Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in mathematics at the LEA level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.60 | -0.18 | -0.02 | 0.04 | 0.09 | 0.25 | 0.87 |
| White–American Indian | -0.49 | -0.49 | -0.07 | 0.08 | 0.25 | 0.71 | 0.71 |
| White–Asian | -1.52 | -0.53 | -0.24 | -0.16 | -0.04 | 0.26 | 0.65 |
| White–Pacific Islander | -1.09 | -0.47 | -0.09 | 0.06 | 0.23 | 0.69 | -0.76 |
| White–Filipino | -0.85 | -0.47 | -0.20 | -0.10 | 0.00 | 0.28 | 0.61 |
| White–Hispanic | -0.83 | -0.24 | 0.00 | 0.08 | 0.16 | 0.41 | 0.73 |
| White–Black | -0.44 | -0.21 | 0.06 | 0.16 | 0.24 | 0.51 | 0.73 |
| White–Two or more races | -1.06 | -0.40 | -0.12 | -0.01 | 0.08 | 0.36 | 0.87 |
| White–Missing (Unknown) | -0.77 | -0.46 | -0.13 | 0.05 | 0.24 | 0.62 | 1.02 |
| Never EL–Current EL | -0.75 | -0.23 | 0.06 | 0.15 | 0.25 | 0.51 | 1.03 |
| Never EL–Former EL | -0.92 | -0.42 | -0.16 | -0.07 | 0.01 | 0.27 | 0.55 |
| Never EL–Ever EL | -0.67 | -0.31 | -0.05 | 0.04 | 0.13 | 0.38 | 0.77 |
| SWD–Not SWD | -1.12 | -0.62 | -0.35 | -0.26 | -0.16 | 0.10 | 0.74 |
| SED–Not SED | -1.15 | -0.41 | -0.20 | -0.13 | -0.05 | 0.16 | 0.60 |
| Homeless–Not homeless | -0.94 | -0.50 | -0.20 | -0.08 | 0.01 | 0.27 | 0.75 |
| Foster youth–Not foster youth | -0.73 | -0.73 | -0.30 | -0.14 | -0.01 | 0.32 | 0.32 |

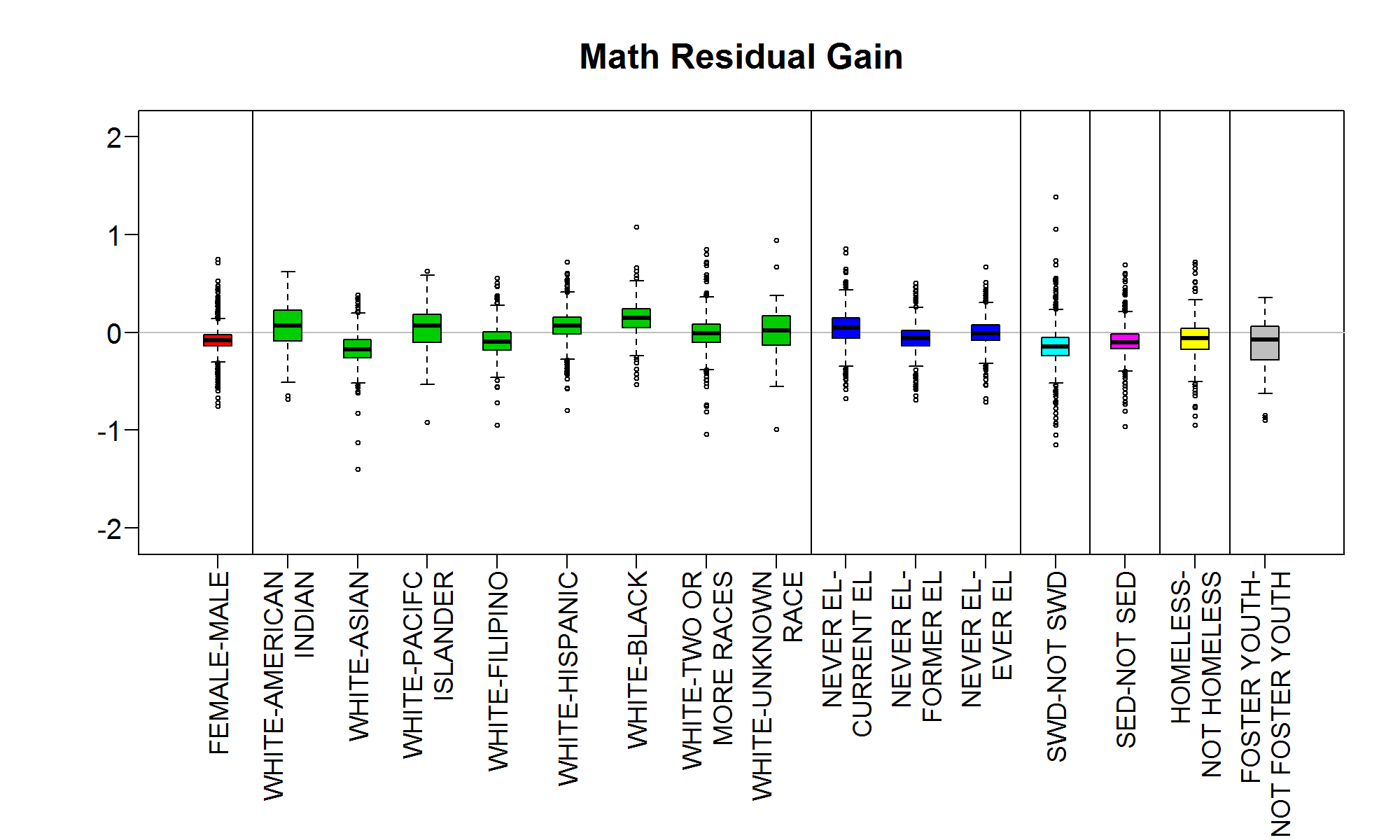


Figure C.3.C. Distributions of group differences (or gaps) for Residual Gain in mathematics at the LEA level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.10. Data Supporting Figure C.3.C: Distributions of group differences (or gaps) for Residual Gain in mathematics at the LEA level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.75 | -0.30 | -0.14 | -0.08 | -0.02 | 0.14 | 0.75 |
| White–American Indian | -0.68 | -0.51 | -0.09 | 0.07 | 0.23 | 0.62 | -0.64 |
| White–Asian | -1.40 | -0.52 | -0.26 | -0.17 | -0.07 | 0.20 | 0.39 |
| White–Pacific Islander | -0.92 | -0.53 | -0.10 | 0.07 | 0.19 | 0.58 | 0.63 |
| White–Filipino | -0.94 | -0.46 | -0.18 | -0.09 | 0.01 | 0.28 | 0.55 |
| White–Hispanic | -0.80 | -0.27 | -0.02 | 0.07 | 0.15 | 0.41 | 0.72 |
| White–Black | -0.53 | -0.24 | 0.05 | 0.15 | 0.24 | 0.53 | 1.08 |
| White–Two or more races | -1.04 | -0.38 | -0.10 | -0.01 | 0.09 | 0.36 | 0.85 |
| White–Missing (Unknown) | -0.99 | -0.55 | -0.13 | 0.02 | 0.17 | 0.38 | 0.95 |
| Never EL–Current EL | -0.67 | -0.35 | -0.06 | 0.05 | 0.15 | 0.44 | 0.86 |
| Never EL–Former EL | -0.69 | -0.34 | -0.14 | -0.06 | 0.02 | 0.26 | 0.51 |
| Never EL–Ever EL | -0.71 | -0.32 | -0.08 | -0.01 | 0.07 | 0.31 | 0.67 |
| SWD–Not SWD | -1.14 | -0.52 | -0.24 | -0.15 | -0.05 | 0.23 | 1.38 |
| SED–Not SED | -0.96 | -0.40 | -0.17 | -0.10 | -0.02 | 0.21 | 0.69 |
| Homeless–Not homeless | -0.94 | -0.50 | -0.18 | -0.06 | 0.04 | 0.34 | 0.72 |
| Foster youth–Not foster youth | -0.90 | -0.63 | -0.28 | -0.08 | 0.07 | 0.36 | -0.85 |

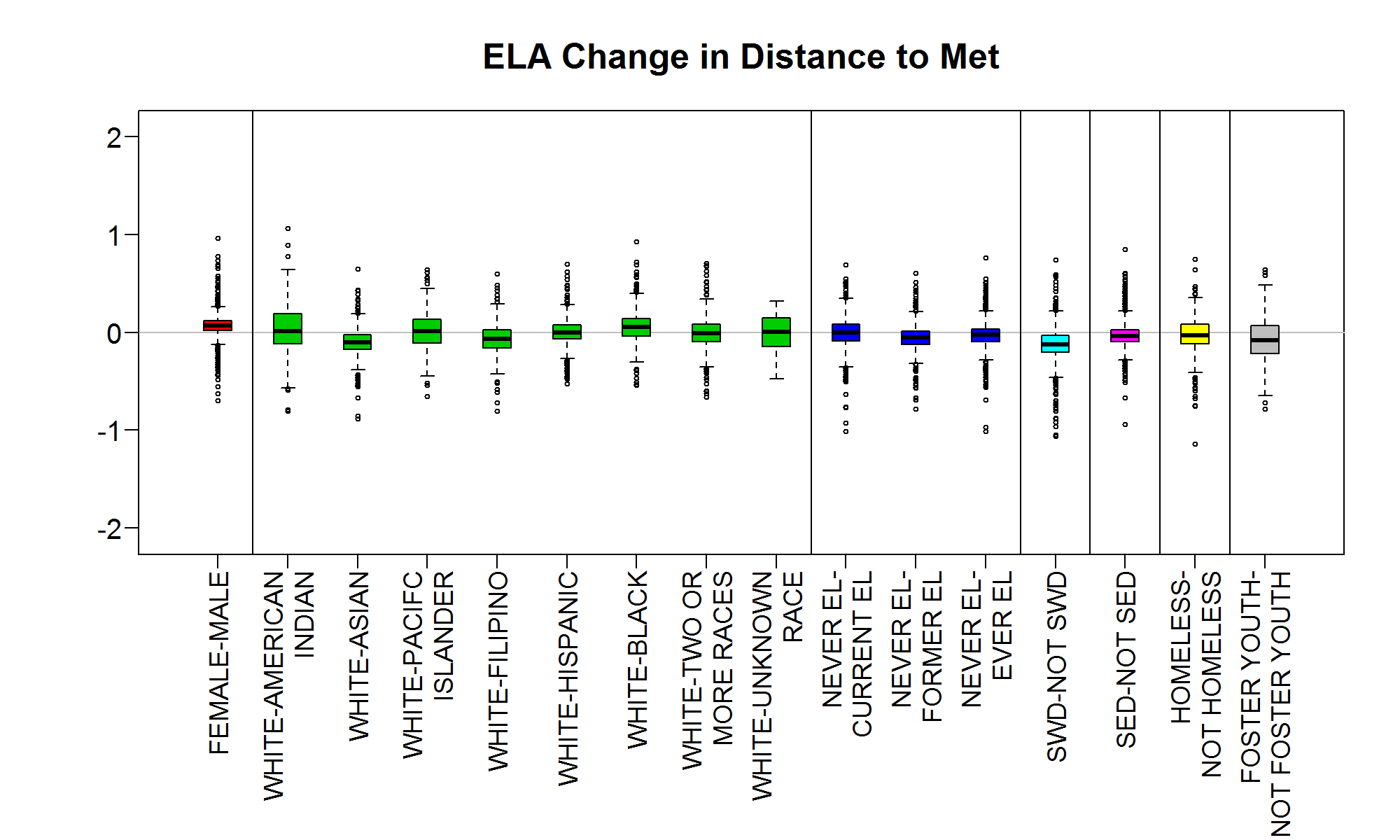


Figure C.4.A. Distributions of group differences (or gaps) for Change in Distance to Met in ELA at the LEA level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.11. Data Supporting Figure C.4.A: Distributions of group differences (or gaps) for Change in Distance to Met in ELA at the LEA level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.70 | -0.12 | 0.02 | 0.07 | 0.12 | 0.27 | 0.97 |
| White–American Indian | -0.81 | -0.57 | -0.12 | 0.02 | 0.19 | 0.64 | 1.06 |
| White–Asian | -0.88 | -0.38 | -0.17 | -0.10 | -0.02 | 0.19 | 0.65 |
| White–Pacific Islander | -0.65 | -0.44 | -0.11 | 0.01 | 0.13 | 0.45 | 0.64 |
| White–Filipino | -0.80 | -0.43 | -0.16 | -0.06 | 0.02 | 0.29 | 0.60 |
| White–Hispanic | -0.52 | -0.27 | -0.07 | 0.00 | 0.08 | 0.29 | 0.70 |
| White–Black | -0.54 | -0.30 | -0.04 | 0.06 | 0.14 | 0.40 | 0.93 |
| White–Two or more races | -0.66 | -0.35 | -0.09 | -0.01 | 0.08 | 0.34 | 0.71 |
| White–Missing (Unknown) | -0.47 | -0.47 | -0.14 | 0.01 | 0.15 | 0.32 | 0.32 |
| Never EL–Current EL | -1.01 | -0.35 | -0.09 | 0.00 | 0.09 | 0.35 | 0.69 |
| Never EL–Former EL | -0.78 | -0.32 | -0.12 | -0.05 | 0.01 | 0.22 | 0.61 |
| Never EL–Ever EL | -1.01 | -0.28 | -0.09 | -0.02 | 0.03 | 0.22 | 0.76 |
| SWD–Not SWD | -1.06 | -0.46 | -0.20 | -0.12 | -0.03 | 0.22 | 0.74 |
| SED–Not SED | -0.94 | -0.28 | -0.10 | -0.04 | 0.03 | 0.22 | 0.85 |
| Homeless–Not homeless | -1.14 | -0.41 | -0.12 | -0.03 | 0.08 | 0.36 | 0.75 |
| Foster youth–Not foster youth | -0.78 | -0.64 | -0.22 | -0.08 | 0.07 | 0.48 | 0.64 |

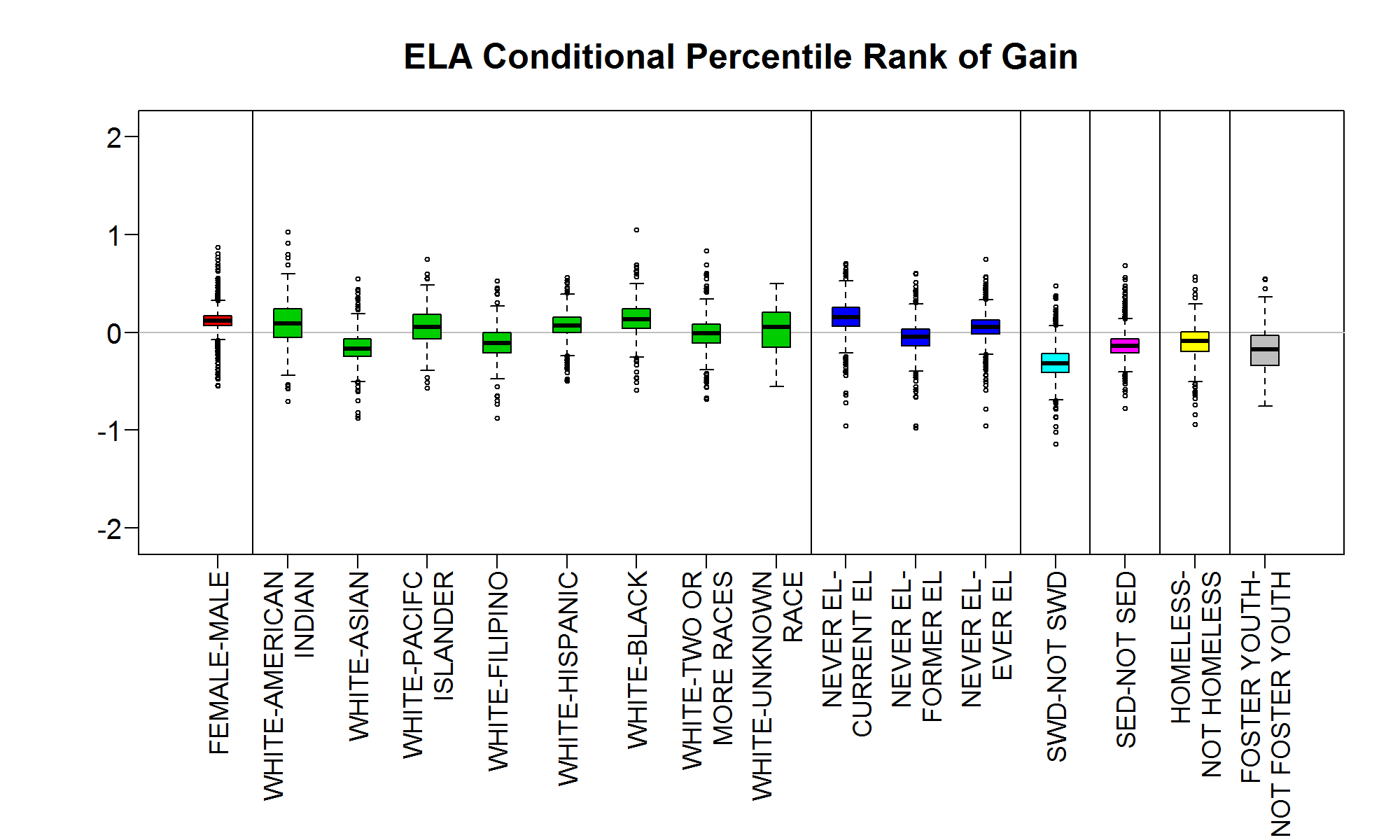


Figure C.4.B. Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in ELA at the LEA level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.12. Data Supporting Figure C.4.B: Distributions of group differences (or gaps) for Conditional Percentile Rank of Gain in ELA at the LEA level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.55 | -0.07 | 0.07 | 0.12 | 0.17 | 0.32 | 0.87 |
| White–American Indian | -0.70 | -0.44 | -0.05 | 0.09 | 0.24 | 0.60 | 1.03 |
| White–Asian | -0.87 | -0.50 | -0.24 | -0.17 | -0.07 | 0.19 | 0.55 |
| White–Pacific Islander | -0.57 | -0.39 | -0.06 | 0.06 | 0.18 | 0.49 | 0.75 |
| White–Filipino | -0.88 | -0.47 | -0.21 | -0.11 | 0.00 | 0.27 | 0.53 |
| White–Hispanic | -0.50 | -0.24 | 0.00 | 0.07 | 0.15 | 0.39 | 0.56 |
| White–Black | -0.59 | -0.25 | 0.04 | 0.14 | 0.24 | 0.50 | 1.05 |
| White–Two or more races | -0.68 | -0.38 | -0.11 | -0.01 | 0.08 | 0.34 | 0.84 |
| White–Missing (Unknown) | -0.55 | -0.55 | -0.15 | 0.06 | 0.20 | 0.50 | 0.50 |
| Never EL–Current EL | -0.96 | -0.21 | 0.06 | 0.16 | 0.25 | 0.53 | 0.71 |
| Never EL–Former EL | -0.97 | -0.40 | -0.14 | -0.05 | 0.03 | 0.29 | 0.61 |
| Never EL–Ever EL | -0.96 | -0.22 | -0.02 | 0.06 | 0.12 | 0.33 | 0.75 |
| SWD–Not SWD | -1.14 | -0.69 | -0.41 | -0.32 | -0.22 | 0.07 | 0.48 |
| SED–Not SED | -0.77 | -0.40 | -0.21 | -0.14 | -0.07 | 0.14 | 0.69 |
| Homeless–Not homeless | -0.94 | -0.50 | -0.20 | -0.09 | 0.01 | 0.29 | 0.57 |
| Foster youth–Not foster youth | 0.45 | -0.75 | -0.34 | -0.17 | -0.03 | 0.36 | 0.55 |

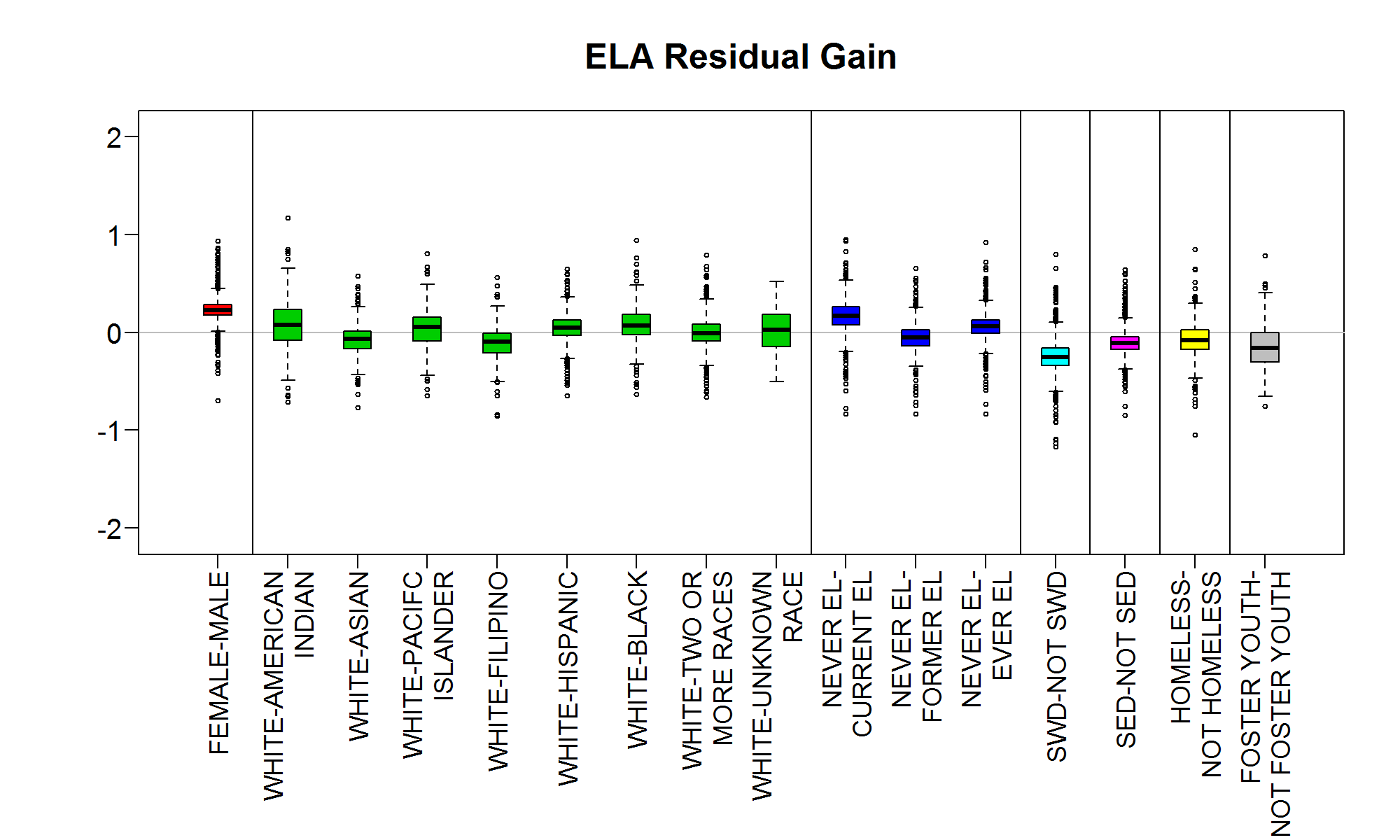


Figure C.4.C. Distributions of group differences (or gaps) for Residual Gain in ELA at the LEA level for the paired student groups listed along the x-axis. Additionally, comparisons of the student groups listed along the x-axis within the same group type are organized within vertical lines and coded with the same color (i.e., all comparisons between White and other races/ethnicities are colored green).

Table C.13. Data Supporting Figure C.4.C: Distributions of group differences (or gaps) for Residual Gain in ELA at the LEA level.

| Paired Student Groups Listed Along the X-Axis | Minimum | Lower Whisker | 25th Quantile | Median | 75th Quantile | Upper Whisker | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female–Male | -0.70 | 0.02 | 0.18 | 0.23 | 0.29 | 0.45 | 0.94 |
| White–American Indian | -0.71 | -0.49 | -0.08 | 0.08 | 0.23 | 0.66 | 1.17 |
| White–Asian | -0.77 | -0.43 | -0.17 | -0.07 | 0.01 | 0.26 | 0.58 |
| White–Pacific Islander | -0.65 | -0.44 | -0.09 | 0.05 | 0.16 | 0.49 | 0.80 |
| White–Filipino | -0.86 | -0.50 | -0.21 | -0.10 | -0.01 | 0.27 | 0.57 |
| White–Hispanic | -0.65 | -0.27 | -0.03 | 0.05 | 0.13 | 0.36 | 0.65 |
| White–Black | -0.63 | -0.33 | -0.03 | 0.07 | 0.18 | 0.49 | 0.94 |
| White–Two or more races | -0.66 | -0.34 | -0.09 | -0.01 | 0.09 | 0.34 | 0.79 |
| White–Missing (Unknown) | -0.50 | -0.50 | -0.15 | 0.02 | 0.19 | 0.52 | 0.52 |
| Never EL–Current EL | -0.83 | -0.20 | 0.08 | 0.17 | 0.26 | 0.53 | 0.95 |
| Never EL–Former EL | -0.83 | -0.35 | -0.13 | -0.05 | 0.03 | 0.26 | 0.66 |
| Never EL–Ever EL | -0.83 | -0.21 | -0.01 | 0.06 | 0.13 | 0.33 | 0.92 |
| SWD–Not SWD | -1.17 | -0.60 | -0.34 | -0.25 | -0.16 | 0.11 | 0.80 |
| SED–Not SED | -0.84 | -0.37 | -0.18 | -0.11 | -0.04 | 0.15 | 0.64 |
| Homeless–Not homeless | -1.04 | -0.47 | -0.17 | -0.08 | 0.03 | 0.30 | 0.85 |
| Foster youth–Not foster youth | -0.76 | -0.65 | -0.30 | -0.16 | 0.00 | 0.41 | 0.78 |

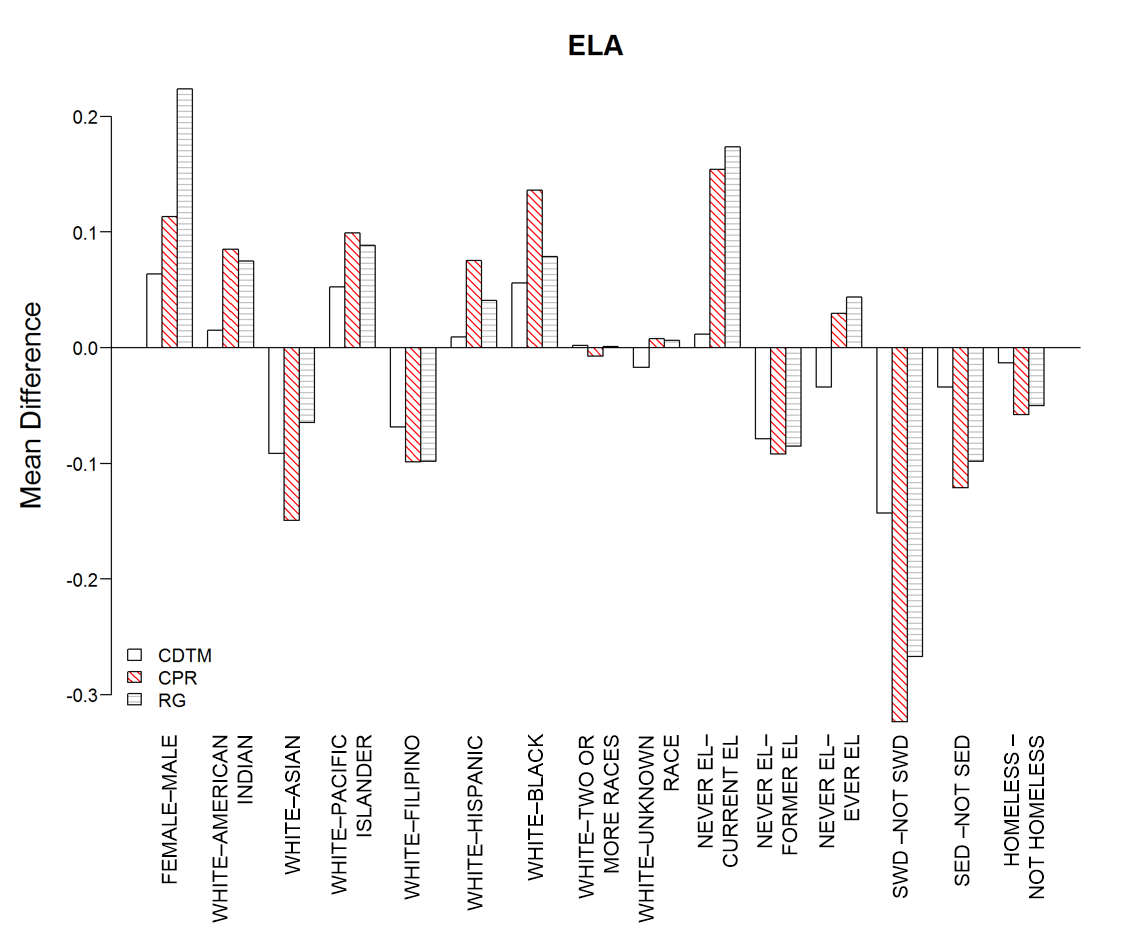


Figure C.5.A. Mean group differences for ELA at the school level for the paired student groups listed along the x‑axis. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table C.14. Data Supporting Figure C.5.A: Mean group differences for ELA at the school level.

| Paired Student Groups Listed Along the X-Axis | CDTM | CPR | RG |
| --- | --- | --- | --- |
| Female–Male | 0.06 | 0.11 | 0.22 |
| White–American Indian | 0.01 | 0.09 | 0.07 |
| White–Asian | -0.09 | -0.15 | -0.06 |
| White–Pacific Islander | 0.05 | 0.10 | 0.09 |
| White–Filipino | -0.07 | -0.10 | -0.10 |
| White–Hispanic | 0.01 | 0.08 | 0.04 |
| White–Black | 0.06 | 0.14 | 0.08 |
| White–Two or more races | 0.00 | -0.01 | 0.00 |
| White–Missing (Unknown) | -0.02 | 0.01 | 0.01 |
| Never EL–Current EL | 0.01 | 0.15 | 0.17 |
| Never EL–Former EL | -0.08 | -0.09 | -0.09 |
| Never EL–Ever EL | -0.03 | 0.03 | 0.04 |
| SWD–Not SWD | -0.14 | -0.32 | -0.27 |
| SED–Not SED | -0.03 | -0.12 | -0.10 |
| Homeless–Not homeless | -0.01 | -0.06 | -0.05 |

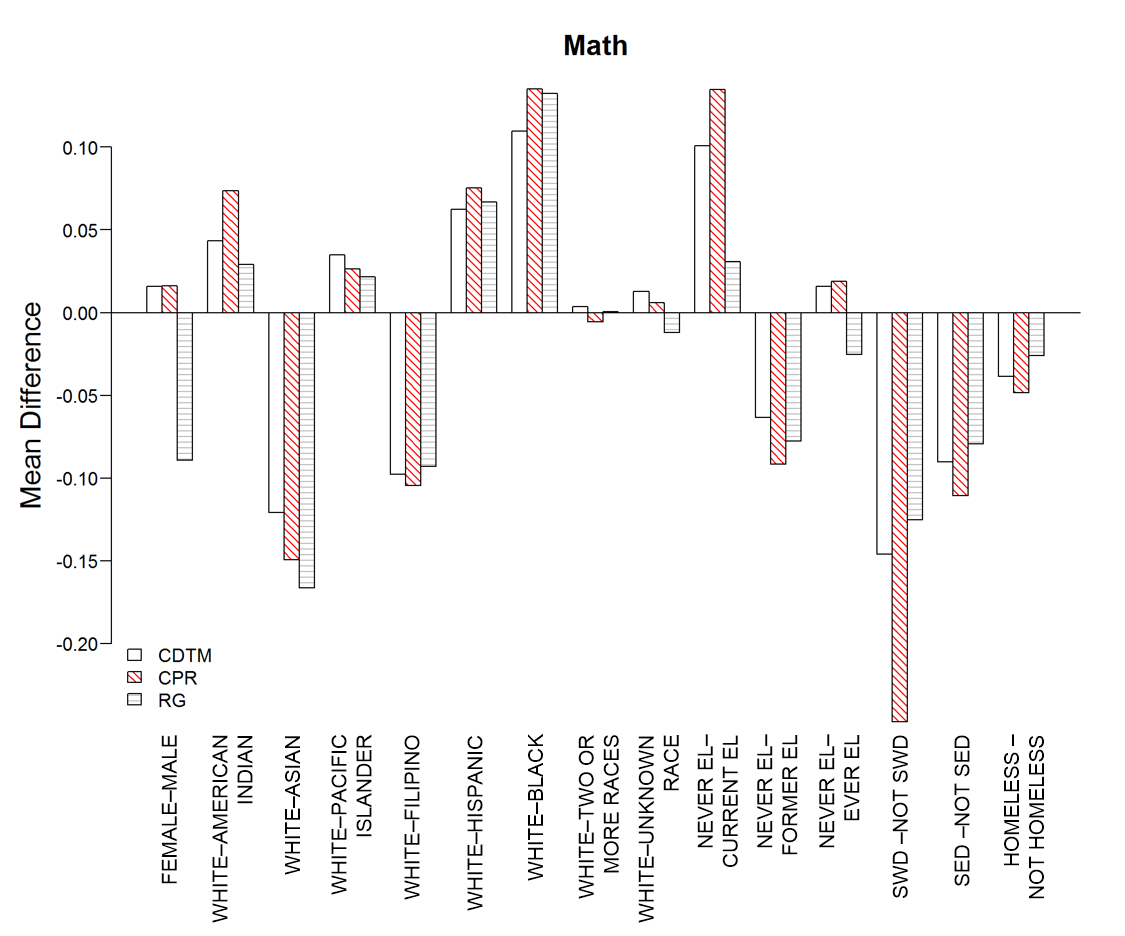


Figure C.5.B. Mean group differences for mathematics at the school level for the paired student groups listed along the x-axis. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table C.15. Data Supporting Figure C.5.A: Mean group differences for mathematics at the school level.

| Paired Student Groups Listed Along the X-Axis | CDTM | CPR | RG |
| --- | --- | --- | --- |
| Female–Male | 0.02 | 0.02 | -0.09 |
| White–American Indian | 0.04 | 0.07 | 0.03 |
| White–Asian | -0.12 | -0.15 | -0.17 |
| White–Pacific Islander | 0.03 | 0.03 | 0.02 |
| White–Filipino | -0.10 | -0.10 | -0.09 |
| White–Hispanic | 0.06 | 0.08 | 0.07 |
| White–Black | 0.11 | 0.14 | 0.13 |
| White–Two or more races | 0.00 | -0.01 | 0.00 |
| White–Missing (Unknown) | 0.01 | 0.01 | -0.01 |
| Never EL–Current EL | 0.10 | 0.13 | 0.03 |
| Never EL–Former EL | -0.06 | -0.09 | -0.08 |
| Never EL–Ever EL | 0.02 | 0.02 | -0.03 |
| SWD–Not SWD | -0.15 | -0.25 | -0.13 |
| SED–Not SED | -0.09 | -0.11 | -0.08 |
| Homeless–Not homeless | -0.04 | -0.05 | -0.03 |

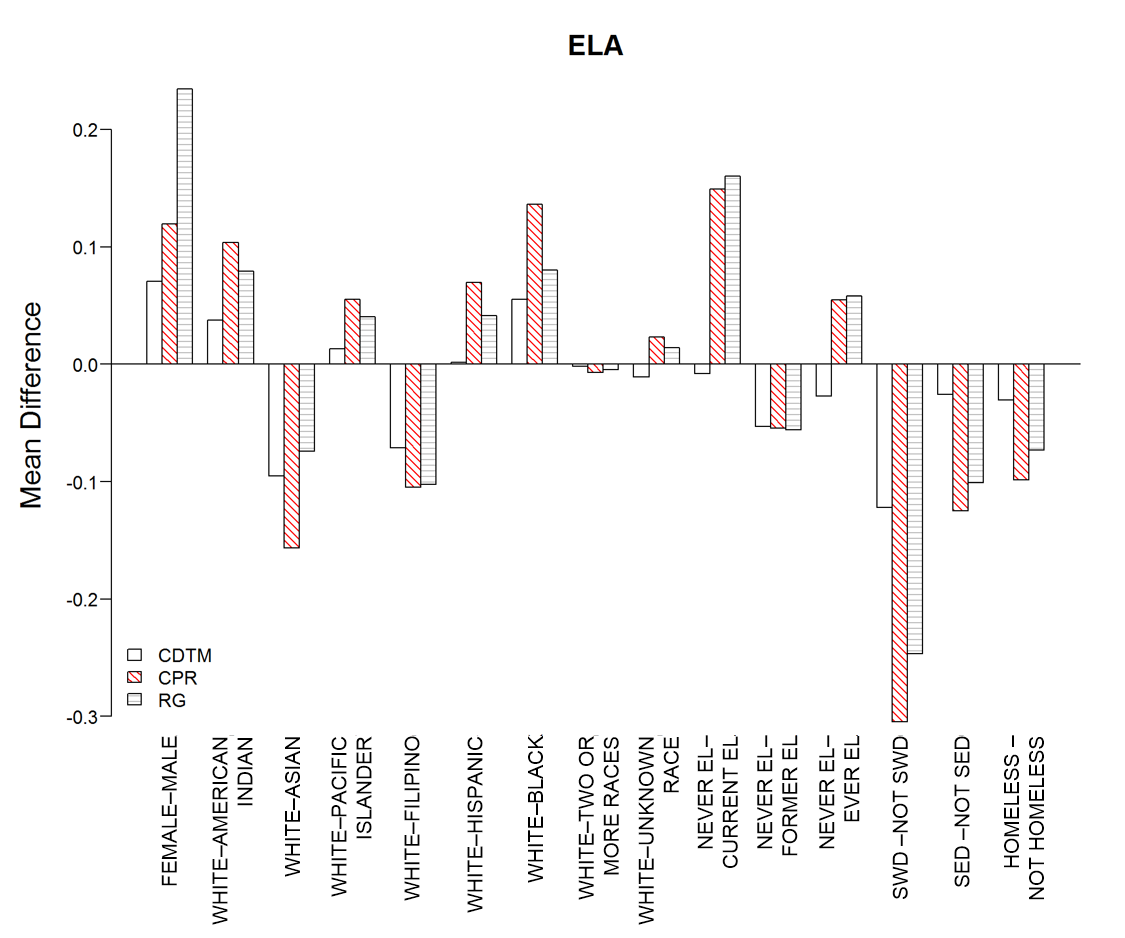


Figure C.6.A. Mean group differences for ELA at the LEA level for the paired student groups listed along the x-axis. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table C.16. Data Supporting Figure C.6.A: Mean group differences for ELA at the LEA level.

| Paired Student Groups Listed Along the X-Axis | CDTM | CPR | RG |
| --- | --- | --- | --- |
| Female–Male | 0.07 | 0.12 | 0.23 |
| White–American Indian | 0.04 | 0.10 | 0.08 |
| White–Asian | -0.10 | -0.16 | -0.07 |
| White–Pacific Islander | 0.01 | 0.06 | 0.04 |
| White–Filipino | -0.07 | -0.10 | -0.10 |
| White–Hispanic | 0.00 | 0.07 | 0.04 |
| White–Black | 0.06 | 0.14 | 0.08 |
| White–Two or more races | 0.00 | -0.01 | 0.00 |
| White–Missing (Unknown) | -0.01 | 0.02 | 0.01 |
| Never EL–Current EL | -0.01 | 0.15 | 0.16 |
| Never EL–Former EL | -0.05 | -0.05 | -0.06 |
| Never EL–Ever EL | -0.03 | 0.05 | 0.06 |
| SWD–Not SWD | -0.12 | -0.30 | -0.25 |
| SED–Not SED | -0.03 | -0.12 | -0.10 |
| Homeless–Not homeless | -0.03 | -0.10 | -0.07 |

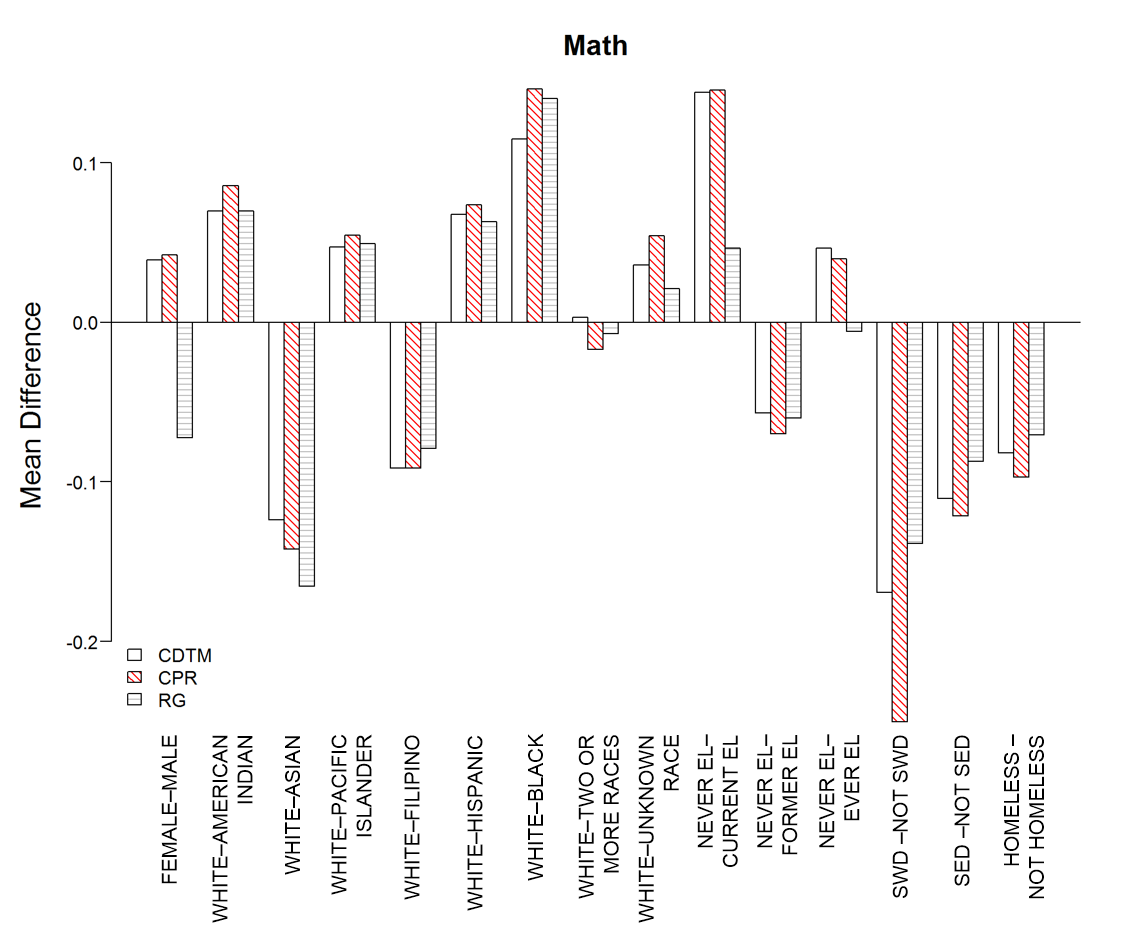


Figure C.6.B. Mean group differences for mathematics at the LEA level for the paired student groups listed along the x-axis. (Note that CDTM = change in distance-to-met, CPR=conditional percentile rank of gain, and RG=residual gain.)

Table C.17. Data Supporting Figure C.6.B: Mean group differences for mathematics at the LEA.

| Paired Student Groups Listed Along the X-Axis | CDTM | CPR | RG |
| --- | --- | --- | --- |
| Female–Male | 0.04 | 0.04 | -0.07 |
| White–American Indian | 0.07 | 0.09 | 0.07 |
| White–Asian | -0.12 | -0.14 | -0.17 |
| White–Pacific Islander | 0.05 | 0.05 | 0.05 |
| White–Filipino | -0.09 | -0.09 | -0.08 |
| White–Hispanic | 0.07 | 0.07 | 0.06 |
| White–Black | 0.11 | 0.15 | 0.14 |
| White–Two or more races | 0.00 | -0.02 | -0.01 |
| White–Missing (Unknown) | 0.04 | 0.05 | 0.02 |
| Never EL–Current EL | 0.14 | 0.15 | 0.05 |
| Never EL–Former EL | -0.06 | -0.07 | -0.06 |
| Never EL–Ever EL | 0.05 | 0.04 | -0.01 |
| SWD–Not SWD | -0.17 | -0.25 | -0.14 |
| SED–Not SED | -0.11 | -0.12 | -0.09 |
| Homeless–Not homeless | -0.08 | -0.10 | -0.07 |

# Appendix D: Description of Standard Error and Reliability Computations

For the purposes of these calculations, it is assumed that the sample of observed students in each grade and school (or local educational agency [LEA]) is a random sample from a hypothetical population of students corresponding to that grade and school (or LEA). Without loss of generality, to provide a concrete example, we describe the procedure that we used to compute the standard errors of the school-level aggregates of the mathematics change in distance-to-met (CDTM) growth measure.

We started with the data consisting of the mathematics CDTM growth measure at the individual student level, the grade level of each student, and the school to which each student is linked. We split this dataset by grade. For each grade, and for each school, we computed the sample variance of the mathematics CDTM growth measures from that grade and school. For school/grade combinations in which there was only a single student with the mathematics CDTM measure, we set the sample variance to the weighted average of the sample variances computed for all schools, where weights were determined by the number of students in each school for the given grade. At this point, for a given grade, we have computed a sample variance of the mathematics CDTM measures for each school.

We then divided this by the number of students in the given grade in each school, which is the usual formula for estimating the variance of a mean from a random sample, that is, . Thus, we have estimated the variance of the grade-level mean mathematics CDTM for each school.

The school-level aggregate mathematics CDTM for each school can be expressed as a weighted average of the corresponding grade-level means, where the weights are the relative proportions of students in each grade. Under the independence assumption previously noted, the variance of this weighted average can be computed with a standard formula—that is, if  equals the mean for grade , then  =  where —from the weights () and the grade-specific variance estimates  already computed. The standard errors are obtained by square root. This procedure was replicated for mathematics and ELA, each growth measure, and at the school and LEA levels.

We then used these variance estimates for the aggregate growth statistics to estimate a reliability for each content area, growth statistic, and level of aggregation. The reliability was computed using the method of moments as follows, again using the mathematics CDTM at the school-level as an example:

1. We restricted to schools with at least 10 students contributing to the mathematics CDTM school-level average.
2. We computed the sample variance of the school aggregate growth measures across schools and called this “V.”
3. We also computed the average across schools of the estimated variance of each school’s aggregate growth measure (as described in previous paragraph) and called this “E.”
4. We then estimated the reliability as



The numerator is an estimate of the signal variance of the aggregate growth statistic across schools—that is, it estimates how much variance there would be across schools in the aggregate growth statistic if each school had an infinite number of students from each grade. The denominator is the total variance of the observed measures, which includes the signal variance as well as the extra variability across schools due to finite sample of students in each school and grade. The ratio of the signal variance to the total variance is commonly used as a measure of reliability.

# Appendix E: Comparisons of Rankings by Growth Measures for School/LEA Student Groups

Table E.1. Spearman rank-order correlations between each pair of growth measures for student group means at the school level

| Student group | Growth Measure 1 | Growth Measure 2 | ELA Correlation | Mathematics Correlation | N |
| --- | --- | --- | --- | --- | --- |
| Female | CDTM | RG | 0.89 | 0.89 | 7,214 |
| Female | CDTM | CPR | 0.88 | 0.91 | 7,214 |
| Female | RG | CPR | 0.93 | 0.96 | 7,214 |
| Male | CDTM | RG | 0.92 | 0.91 | 7,205 |
| Male | CDTM | CPR | 0.91 | 0.90 | 7,205 |
| Male | RG | CPR | 0.93 | 0.96 | 7,205 |
| American Indian | CDTM | RG | 0.92 | 0.94 | 138 |
| American Indian | CDTM | CPR | 0.92 | 0.92 | 138 |
| American Indian | RG | CPR | 0.92 | 0.92 | 138 |
| Asian | CDTM | RG | 0.89 | 0.86 | 2,967 |
| Asian | CDTM | CPR | 0.86 | 0.83 | 2,967 |
| Asian | RG | CPR | 0.88 | 0.93 | 2,967 |
| Pacific Islander | CDTM | RG | 0.89 | 0.92 | 154 |
| Pacific Islander | CDTM | CPR | 0.92 | 0.91 | 154 |
| Pacific Islander | RG | CPR | 0.92 | 0.91 | 154 |
| Filipino | CDTM | RG | 0.90 | 0.86 | 1,258 |
| Filipino | CDTM | CPR | 0.92 | 0.87 | 1,258 |
| Filipino | RG | CPR | 0.92 | 0.94 | 1,258 |
| Hispanic | CDTM | RG | 0.91 | 0.89 | 6,919 |
| Hispanic | CDTM | CPR | 0.93 | 0.91 | 6,919 |
| Hispanic | RG | CPR | 0.95 | 0.96 | 6,919 |
| Black | CDTM | RG | 0.92 | 0.90 | 2,569 |
| Black | CDTM | CPR | 0.92 | 0.91 | 2,569 |
| Black | RG | CPR | 0.94 | 0.93 | 2,569 |
| White | CDTM | RG | 0.91 | 0.88 | 5,119 |
| White | CDTM | CPR | 0.91 | 0.88 | 5,119 |
| White | RG | CPR | 0.93 | 0.95 | 5,119 |
| Two or more races | CDTM | RG | 0.91 | 0.87 | 2,061 |
| Two or more races | CDTM | CPR | 0.91 | 0.86 | 2,061 |
| Two or more races | RG | CPR | 0.92 | 0.94 | 2,061 |
| Missing (Unknown race/ethnicity) | CDTM | RG | 0.91 | 0.88 | 74 |
| Missing (Unknown race/ethnicity) | CDTM | CPR | 0.89 | 0.90 | 74 |
| Missing (Unknown race/ethnicity) | RG | CPR | 0.91 | 0.91 | 74 |
| Ever EL | CDTM | RG | 0.92 | 0.91 | 6,475 |
| Ever EL | CDTM | CPR | 0.93 | 0.91 | 6,475 |
| Ever EL | RG | CPR | 0.93 | 0.96 | 6,475 |
| Former EL | CDTM | RG | 0.91 | 0.89 | 5,716 |
| Former EL | CDTM | CPR | 0.93 | 0.90 | 5,716 |
| Former EL | RG | CPR | 0.92 | 0.96 | 5,716 |
| Never EL | CDTM | RG | 0.91 | 0.90 | 7,277 |
| Never EL | CDTM | CPR | 0.90 | 0.91 | 7,277 |
| Never EL | RG | CPR | 0.93 | 0.96 | 7,277 |
| Current EL | CDTM | RG | 0.93 | 0.93 | 5,702 |
| Current EL | CDTM | CPR | 0.95 | 0.90 | 5,702 |
| Current EL | RG | CPR | 0.93 | 0.95 | 5,702 |
| SWD | CDTM | RG | 0.93 | 0.92 | 6,479 |
| SWD | CDTM | CPR | 0.89 | 0.86 | 6,479 |
| SWD | RG | CPR | 0.90 | 0.92 | 6,479 |
| Not SWD | CDTM | RG | 0.90 | 0.89 | 7,311 |
| Not SWD | CDTM | CPR | 0.90 | 0.90 | 7,311 |
| Not SWD | RG | CPR | 0.93 | 0.97 | 7,311 |
| SED | CDTM | RG | 0.92 | 0.90 | 7,029 |
| SED | CDTM | CPR | 0.94 | 0.91 | 7,029 |
| SED | RG | CPR | 0.95 | 0.96 | 7,029 |
| Not SED | CDTM | RG | 0.91 | 0.88 | 6,209 |
| Not SED | CDTM | CPR | 0.91 | 0.88 | 6,209 |
| Not SED | RG | CPR | 0.91 | 0.96 | 6,209 |
| Homeless | CDTM | RG | 0.92 | 0.90 | 1,640 |
| Homeless | CDTM | CPR | 0.94 | 0.90 | 1,640 |
| Homeless | RG | CPR | 0.93 | 0.93 | 1,640 |
| Not homeless | CDTM | RG | 0.90 | 0.89 | 7,322 |
| Not homeless | CDTM | CPR | 0.89 | 0.90 | 7,322 |
| Not homeless | RG | CPR | 0.93 | 0.97 | 7,322 |
| Foster youth | CDTM | RG | 1.00 | -1.00 | 2 |
| Foster youth | CDTM | CPR | 1.00 | 1.00 | 2 |
| Foster youth | RG | CPR | 1.00 | -1.00 | 2 |
| Not foster youth | CDTM | RG | 0.90 | 0.89 | 7,329 |
| Not foster youth | CDTM | CPR | 0.89 | 0.90 | 7,329 |
| Not foster youth | RG | CPR | 0.93 | 0.97 | 7,329 |

Table E.2. Spearman rank-order correlations between each pair of growth measures for student group means at the LEA level

| Student group | Growth Measure 1 | Growth Measure 2 | ELA Correlation | Mathematics Correlation | N |
| --- | --- | --- | --- | --- | --- |
| Female | CDTM | RG | 0.88 | 0.95 | 806 |
| Female | CDTM | CPR | 0.85 | 0.97 | 806 |
| Female | RG | CPR | 0.91 | 0.94 | 806 |
| Male | CDTM | RG | 0.92 | 0.96 | 807 |
| Male | CDTM | CPR | 0.89 | 0.96 | 807 |
| Male | RG | CPR | 0.91 | 0.95 | 807 |
| American Indian | CDTM | RG | 0.92 | 0.95 | 254 |
| American Indian | CDTM | CPR | 0.93 | 0.93 | 254 |
| American Indian | RG | CPR | 0.91 | 0.92 | 254 |
| Asian | CDTM | RG | 0.89 | 0.94 | 409 |
| Asian | CDTM | CPR | 0.77 | 0.95 | 409 |
| Asian | RG | CPR | 0.81 | 0.94 | 409 |
| Pacific Islander | CDTM | RG | 0.88 | 0.94 | 195 |
| Pacific Islander | CDTM | CPR | 0.90 | 0.93 | 195 |
| Pacific Islander | RG | CPR | 0.89 | 0.89 | 195 |
| Filipino | CDTM | RG | 0.89 | 0.93 | 315 |
| Filipino | CDTM | CPR | 0.92 | 0.95 | 315 |
| Filipino | RG | CPR | 0.89 | 0.92 | 315 |
| Hispanic | CDTM | RG | 0.92 | 0.95 | 748 |
| Hispanic | CDTM | CPR | 0.92 | 0.96 | 748 |
| Hispanic | RG | CPR | 0.94 | 0.94 | 748 |
| Black | CDTM | RG | 0.91 | 0.95 | 365 |
| Black | CDTM | CPR | 0.93 | 0.94 | 365 |
| Black | RG | CPR | 0.93 | 0.92 | 365 |
| White | CDTM | RG | 0.89 | 0.95 | 757 |
| White | CDTM | CPR | 0.89 | 0.96 | 757 |
| White | RG | CPR | 0.92 | 0.94 | 757 |
| Two or more races | CDTM | RG | 0.91 | 0.95 | 446 |
| Two or more races | CDTM | CPR | 0.88 | 0.95 | 446 |
| Two or more races | RG | CPR | 0.90 | 0.95 | 446 |
| Missing (Unknown race/ethnicity) | CDTM | RG | 0.93 | 0.94 | 70 |
| Missing (Unknown race/ethnicity) | CDTM | CPR | 0.92 | 0.94 | 70 |
| Missing (Unknown race/ethnicity) | RG | CPR | 0.91 | 0.91 | 70 |
| Ever EL | CDTM | RG | 0.91 | 0.96 | 675 |
| Ever EL | CDTM | CPR | 0.92 | 0.96 | 675 |
| Ever EL | RG | CPR | 0.90 | 0.95 | 675 |
| Former EL | CDTM | RG | 0.90 | 0.94 | 595 |
| Former EL | CDTM | CPR | 0.91 | 0.96 | 595 |
| Former EL | RG | CPR | 0.88 | 0.94 | 595 |
| Never EL | CDTM | RG | 0.90 | 0.95 | 822 |
| Never EL | CDTM | CPR | 0.88 | 0.97 | 822 |
| Never EL | RG | CPR | 0.91 | 0.95 | 822 |
| Current EL | CDTM | RG | 0.91 | 0.95 | 625 |
| Current EL | CDTM | CPR | 0.93 | 0.94 | 625 |
| Current EL | RG | CPR | 0.91 | 0.94 | 625 |
| SWD | CDTM | RG | 0.93 | 0.95 | 690 |
| SWD | CDTM | CPR | 0.86 | 0.87 | 690 |
| SWD | RG | CPR | 0.90 | 0.89 | 690 |
| Not SWD | CDTM | RG | 0.89 | 0.96 | 825 |
| Not SWD | CDTM | CPR | 0.86 | 0.97 | 825 |
| Not SWD | RG | CPR | 0.91 | 0.95 | 825 |
| SED | CDTM | RG | 0.92 | 0.95 | 790 |
| SED | CDTM | CPR | 0.95 | 0.96 | 790 |
| SED | RG | CPR | 0.94 | 0.95 | 790 |
| Not SED | CDTM | RG | 0.88 | 0.94 | 756 |
| Not SED | CDTM | CPR | 0.87 | 0.97 | 756 |
| Not SED | RG | CPR | 0.87 | 0.95 | 756 |
| Homeless | CDTM | RG | 0.92 | 0.93 | 352 |
| Homeless | CDTM | CPR | 0.93 | 0.92 | 352 |
| Homeless | RG | CPR | 0.93 | 0.92 | 352 |
| Not homeless | CDTM | RG | 0.89 | 0.96 | 828 |
| Not homeless | CDTM | CPR | 0.86 | 0.97 | 828 |
| Not homeless | RG | CPR | 0.91 | 0.96 | 828 |
| Foster youth | CDTM | RG | 0.91 | 0.95 | 112 |
| Foster youth | CDTM | CPR | 0.92 | 0.88 | 112 |
| Foster youth | RG | CPR | 0.87 | 0.86 | 112 |
| Not foster youth | CDTM | RG | 0.89 | 0.96 | 828 |
| Not foster youth | CDTM | CPR | 0.87 | 0.97 | 828 |
| Not foster youth | RG | CPR | 0.91 | 0.95 | 828 |

1. The statuses of Never EL, Former EL, and Current EL are mutually exclusive, but the status of Ever EL equals the status of Current EL or Former EL. These statuses were determined on the basis of CDE variables EL4 and RFEP via the following mappings:

   * **Never EL** is defined as all students with EL4 equal to 0 (which requires the students to have RFEP equal to 0).
   * **Former EL** is defined as classified EL4 equal 1 and RFEP equal to 1.
   * **Current EL** is defined as classified EF4 equal to 1 but RFEP equal to 0.
   * **Ever EL** is defined as all students classified as EL4 equal to 1 regardless of the value of RFEP.

   [↑](#footnote-ref-2)
2. The Economically Disadvantaged status was defined by the CDE variable SED. Students with SED equal to 1 are classified as socioeconomically disadvantaged and students with SED equal to 0 are classified as not socioeconomically disadvantaged. [↑](#footnote-ref-3)
3. All test scores designated with the LOSS condition code had scale scores equal to LOSS for the student’s grade and were excluded from the sample. However, there were several thousand scores across grades, years, and content areas that also were equal to the LOSS but not designated with the LOSS condition code. Per CDE instruction, only scores designated with the LOSS condition code were treated as missing, and all other scores equal to the LOSS were retained in the analyses. [↑](#footnote-ref-4)