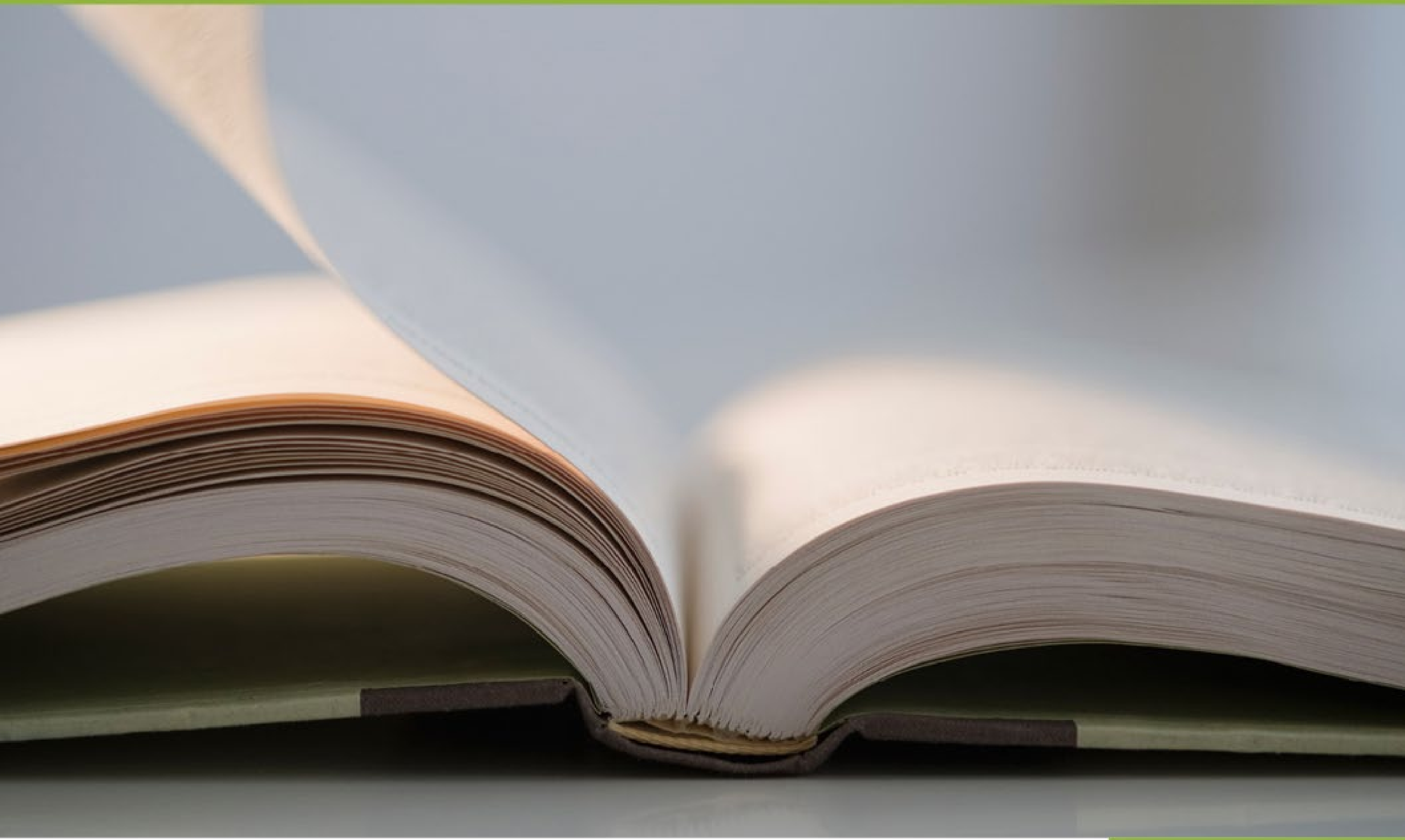
SAS® EVAAS

Research on Missing 2019-20 Data

Prepared for Ohio Department of Education



Contents

Executive Summary

In spring 2020, the COVID-19 pandemic required schools to close early and cancel statewide summative assessments. As a result, scores are not available for the Ohio State Tests (OSTs) based on the 2019-20 school year, and it is not possible to measure growth from the 2018-19 to the 2019-20 school years. The missing test scores will also impact growth measures for the 2020-21 school year.

At the request of the Ohio Department of Education (ODE), SAS investigated the potential impact of missing 2019-20 test scores on the 2020-21 growth measures through a variety of simulations. In essence, these simulations used previous years' test data to compare growth measures with and without a year of missing test scores for different models. These comparisons can provide useful insight as to how the missing 2019-20 test scores could impact the 2020-21 growth measures and assist ODE with its policy decisions regarding this situation.

To summarize findings from simulations with and without a missing year of data, the following outcomes were observed:

- In the gain model, results with a missing year of data were positively correlated with actual results without a missing year of data.
- Furthermore, in the gain model, results with a missing year of data were very closely aligned with actual two-year gains without a missing year of data. This means that gain model results with a missing year of data are most accurately interpreted as a two-year gain.
- That said, two-year growth measures with a missing year of data for the gain model have a different interpretation since students' academic experiences can include districts, schools, and/or teachers from the previous year where data is missing.
- In the predictive model, results with a missing year of data were closely aligned with actual results. In contrast to the gain model, the predictive model is typically used for assessments that are not administered in consecutive grade levels. As a result, although fewer prior scores are available to be used as predictors with a missing year of data, the interpretation is more similar to a typical year than it is for the gain model.

The technical report and appendix include details about the analysis and results.

It is important to note that these simulations are focused solely on the missing year of data and do not estimate the pandemic's impact on student learning in districts, schools, and classrooms. The purpose of this report is to enhance ODE's understanding of how growth models might be impacted by a missing year of assessment data.

In terms of what these findings mean for educators and administrators, the key takeaway is that 2020-21 growth measures are comparable to a two-year measure for the gain model and to a one-year measure for the predictive model. While the interpretation varies in 2020-21 reporting compared to previous years, these measures are still a useful resource for educators, administrators, policymakers and other stakeholders. They can provide insights and answer questions such as the following:

- How did we help our students grow in comparison to students in other schools and districts across the state during these unusual times?
- Did some student populations have more success than others? How did we contribute to that success, and how can we apply what we learned and apply it to future instructional programming?

- Did some student populations have less success? What factors contributed to those results, and how can we help students recover from any incomplete learning?

SAS will work with ODE to provide additional resources for educators and other stakeholders based on the 2020-21 reporting to assist with interpretation and context.

1 Introduction

In spring 2020, the COVID-19 pandemic required schools to close early and cancel statewide summative assessments. As a result, scores are not available for the Ohio State Tests (OSTs) based on the 2019-20 school year, and it is not possible to measure growth from the 2018-19 to the 2019-20 school years. The missing test scores will also impact growth measures for the 2020-21 school year.

Without 2019-20 test scores, it was impossible to provide a growth measure for the 2019-20 school year as there was no current year achievement metric available to use. However, the growth models used in Ohio's value-added reporting can use multiple years of prior test scores to determine students' prior achievement levels. As a result, there are more possibilities for the 2020-21 reporting, which has a current year achievement metric.

This analysis investigates the potential impact of the 2019-20 missing data on results from the gain and predictive models. Each model has a different process for measuring growth and accounting for the missing year of data. A brief description of each model is available in the appendix, and more details are available in the Statistical Models and Business Rules document.

To replicate a variety of scenarios, these models use data for the years 2018-19 and prior. Gain and predictive models were used for subjects and grades in which they are typically used. The analysis uses the model variations below:

- **2019 Gain Model:** 2018-19 growth measures from the gain model *with* 2017-18 test scores
- **2019 Gain Model without 2018 Data:** 2018-19 growth measures from the gain model *without* 2017-18 test scores
- **2019 Gain Model (Two-Year):** sum of single-year standard growth measures from the gain model for 2017-18 and 2018-19 *with* 2017-18 test scores
- **2019 Predictive Model:** 2018-19 growth measures from the predictive model *with* 2017-18 test scores
- **2019 Predictive Model without 2018 Data:** 2018-19 growth measures from the predictive model *without* 2017-18 test scores

The comparisons using these models include:

- **2019 Gain Model versus 2019 Gain Model without 2018 Data:** 2018-19 growth measures from the gain model *with* 2017-18 test scores compared to 2018-19 growth measures from the gain model *without* 2017-18 test scores
- **2019 Gain Model (Two-Year) versus 2019 Gain Model without 2018 Data:** sum of single-year standard growth measures from the gain model for 2017-18 and 2018-19 *with* 2017-18 test scores compared to 2018-19 growth measures from the gain model *without* 2017-18 test scores
- **2019 Predictive Model versus 2019 Predictive Model without 2018 Data:** 2018-19 growth measures from the predictive model *with* 2017-18 test scores compared to 2018-19 growth measures from the predictive model *without* 2017-18 test scores

In this way, ODE can assess the similarities and differences between modeling approaches with and without a particular year of test scores. The results themselves are presented as correlations and scatterplots graphing the district/school/teacher growth index by subject/grade from each model in the comparison. As a reminder, the EVAAS growth measure divided by its standard error provides the *growth index*. This is a standardized statistical value related to the evidence that students' growth is decidedly above or below the growth expectation.

It is important to note that these simulations are focused solely on the missing year of data and do not estimate the pandemic's impact on student learning in districts, schools, and classrooms. Although the results enhance ODE's understanding of how growth models might be impacted by a missing year of assessment data, more analysis is possible when the 2020-21 assessment data becomes available, which might bring forth additional considerations.

2 2019 Gain Model versus 2019 Gain Model without 2018 Data

2.1 Overview

This section compares growth from the 2019 Gain Model to growth from the 2019 Gain Model without 2018 Data. The former model could be considered the standard or typical model for EVAAS reporting in the consecutive grade-given tests with no missing years of data. The latter also uses the standard model but the immediate prior year of test scores is missing. The point of this comparison is to illustrate how removing the immediate prior year of test scores can impact the results of the standard model and understand how the missing 2019-20 school year's data could impact the 2020-21 growth reporting.

Table 1: Overview of Two Models in Comparison

Comparison	2019 Gain Model	2019 Gain Model without 2018 Data
Model approach	Gain Model	Gain Model
One-year growth measure	Yes	No
Gains measure growth from	2017-18 to 2018-19	2016-17 to 2018-19
Inclusion of 2017-18 test scores	Yes	No

For this comparison, correlations and scatterplots can provide insight into the relationship between the two models.

The correlation reports the strength of the relationship between variables with +1 indicating a perfect positive relationship (positive meaning that when one variable changes, the other variable changes in a similar way) and -1 indicating a perfect negative relationship (meaning that when one variable changes, the other variable changes in the opposite direction). Although a precise definition varies, a typical interpretation of the correlation is that a weak relationship is between 0.10 and 0.30, a moderate relationship is between 0.30 and 0.50, and a strong relationship is above 0.50 (Source: Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates). To summarize the results, the district, school, and teacher analyses provide a strong correlation in growth measures between 2019 Gain Model and 2019 Gain Model without 2018 Data.

- The district correlation ranges from 0.48 to 0.62 for individual grades in ELA and from 0.54 to 0.71 for individual grades in Math.
- The school correlation ranges from 0.49 to 0.63 for individual grades in ELA and from 0.55 to 0.71 for individual grades in Math.
- The teacher correlations range from 0.69 to 0.81 for individual grades in ELA and from 0.68 to 0.81 for individual grades in Math.

Correlations for individual subjects and grades are available in the lower right corner of the following scatterplots, which plot the district, school, and teacher growth index for individual subject/grades for ELA and Math. Results for 2019 Gain Model (with 2018 test scores) are on the Y axis, and results for 2019 Gain Model without 2018 Data are on the X axis. The different colors in the graph do not have an assigned meaning; they are simply there to help interpret the distribution more easily.

Even with these strong correlations, it is clear from the tables and scatterplots that the models are not substitutions or close approximations for each other. This is not surprising; one model was based on a one-year gain while the other was based on a two-year gain due to the missing year of data.

2.2 District, School, and Teacher Results

Figure 1: 2019 Gain Model versus 2019 Gain Model without 2018 Data for District-Level Growth Indices in Individual Grades for ELA

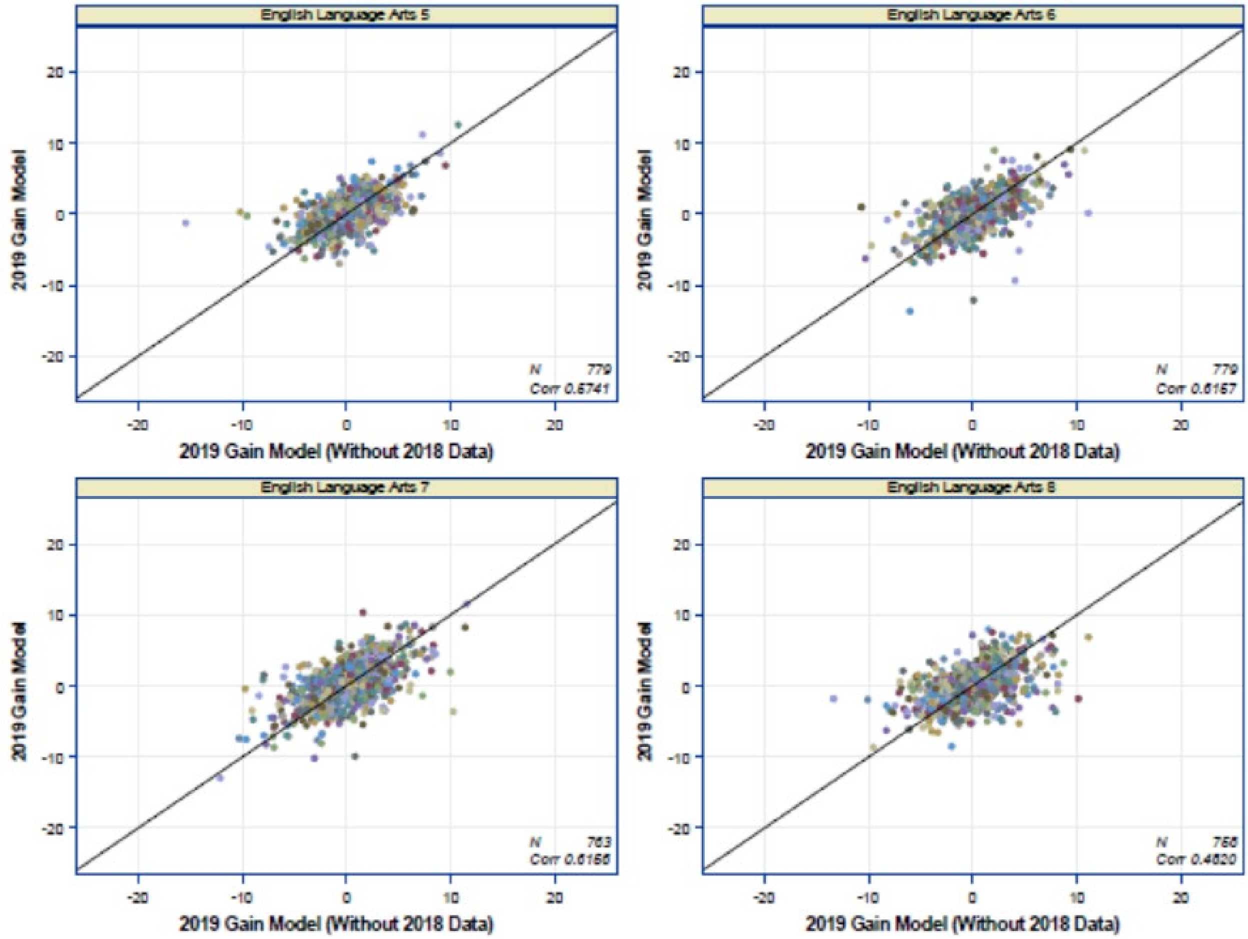


Figure 2: 2019 Gain Model versus 2019 Gain Model without 2018 Data for District-Level Growth Indices in Individual Grades for Math

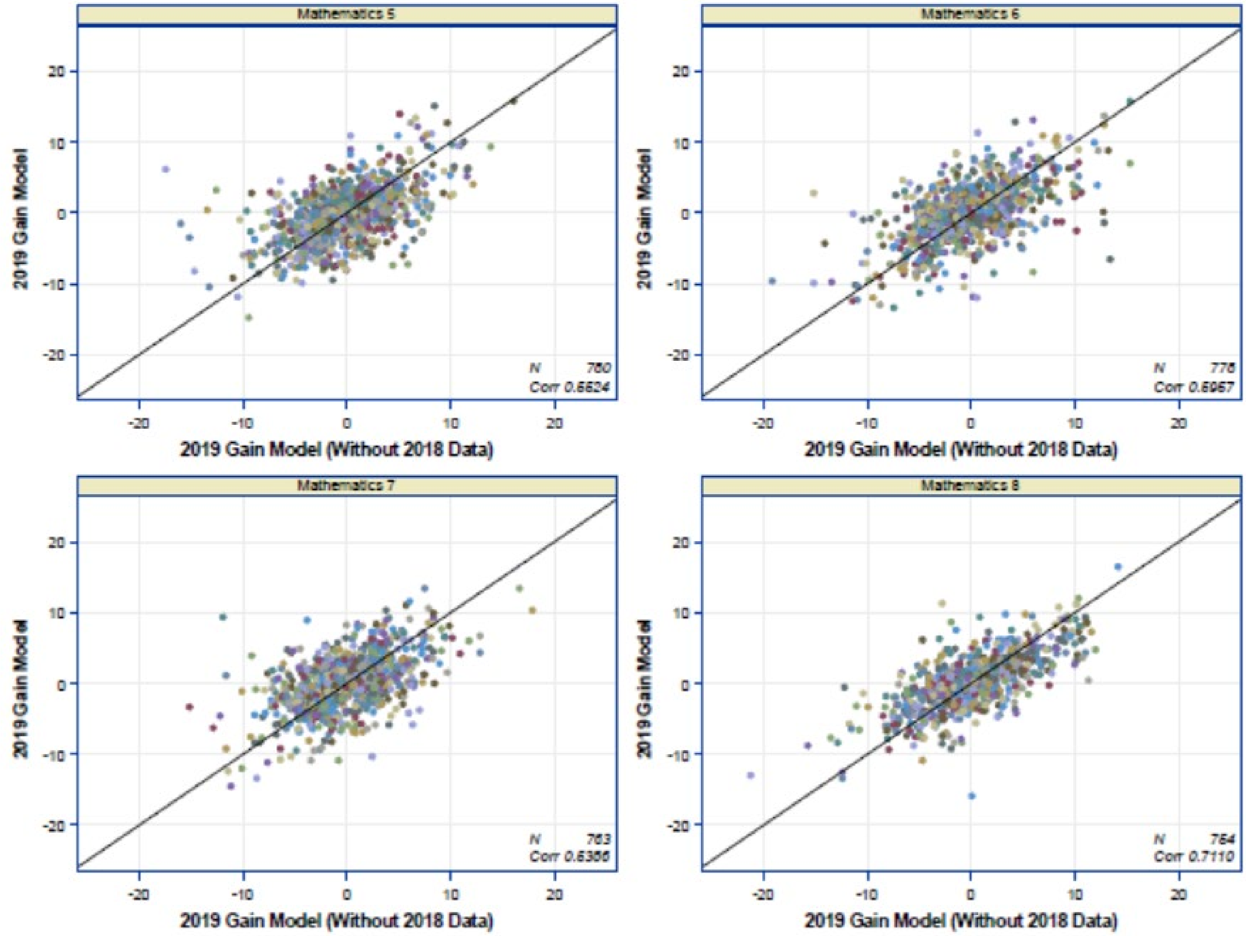


Figure 3: 2019 Gain Model versus 2019 Gain Model without 2018 Data for School-Level Growth Indices in Individual Grades for ELA

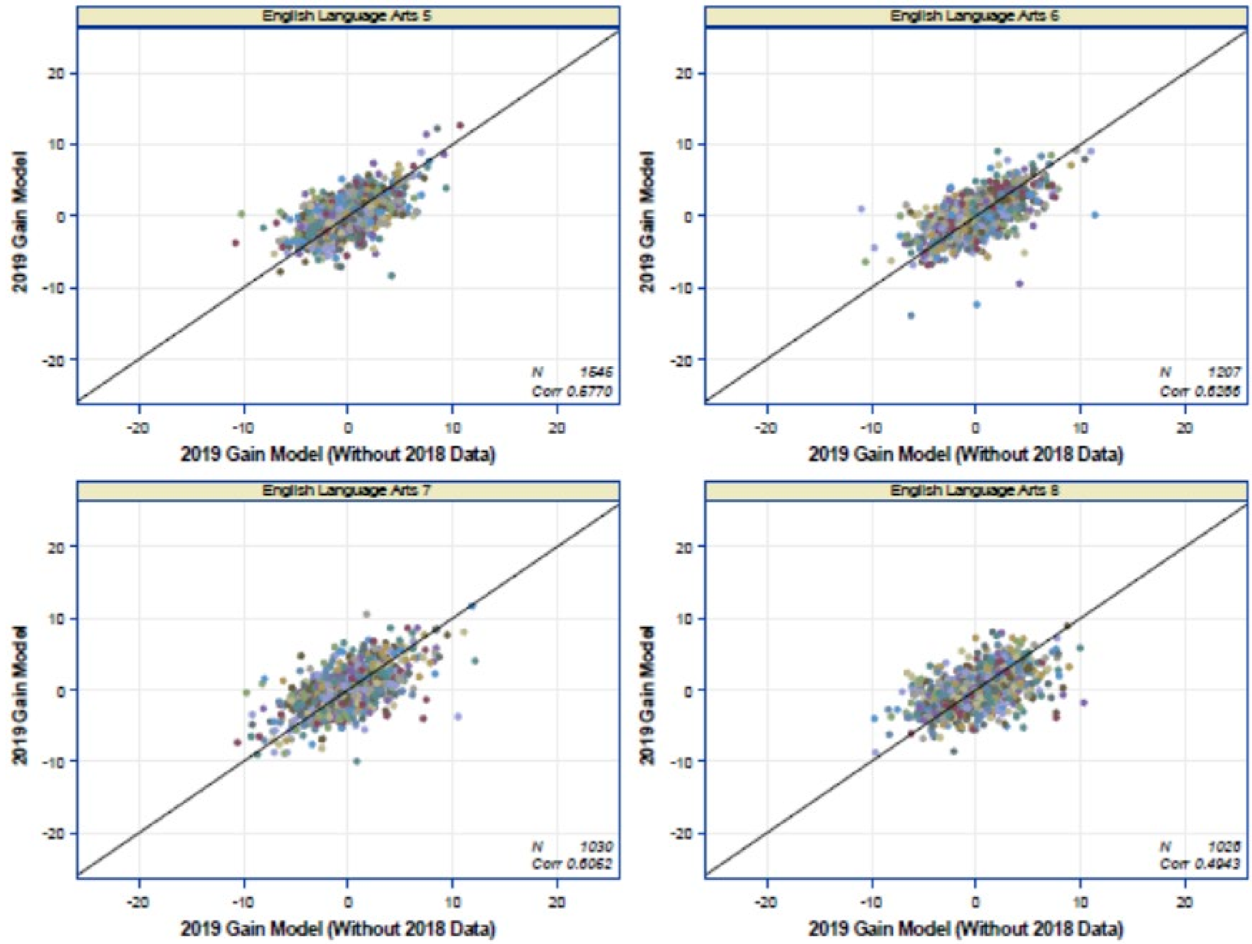


Figure 4: 2019 Gain Model versus 2019 Gain Model without 2018 Data for School-Level Growth Indices in Individual Grades for Math

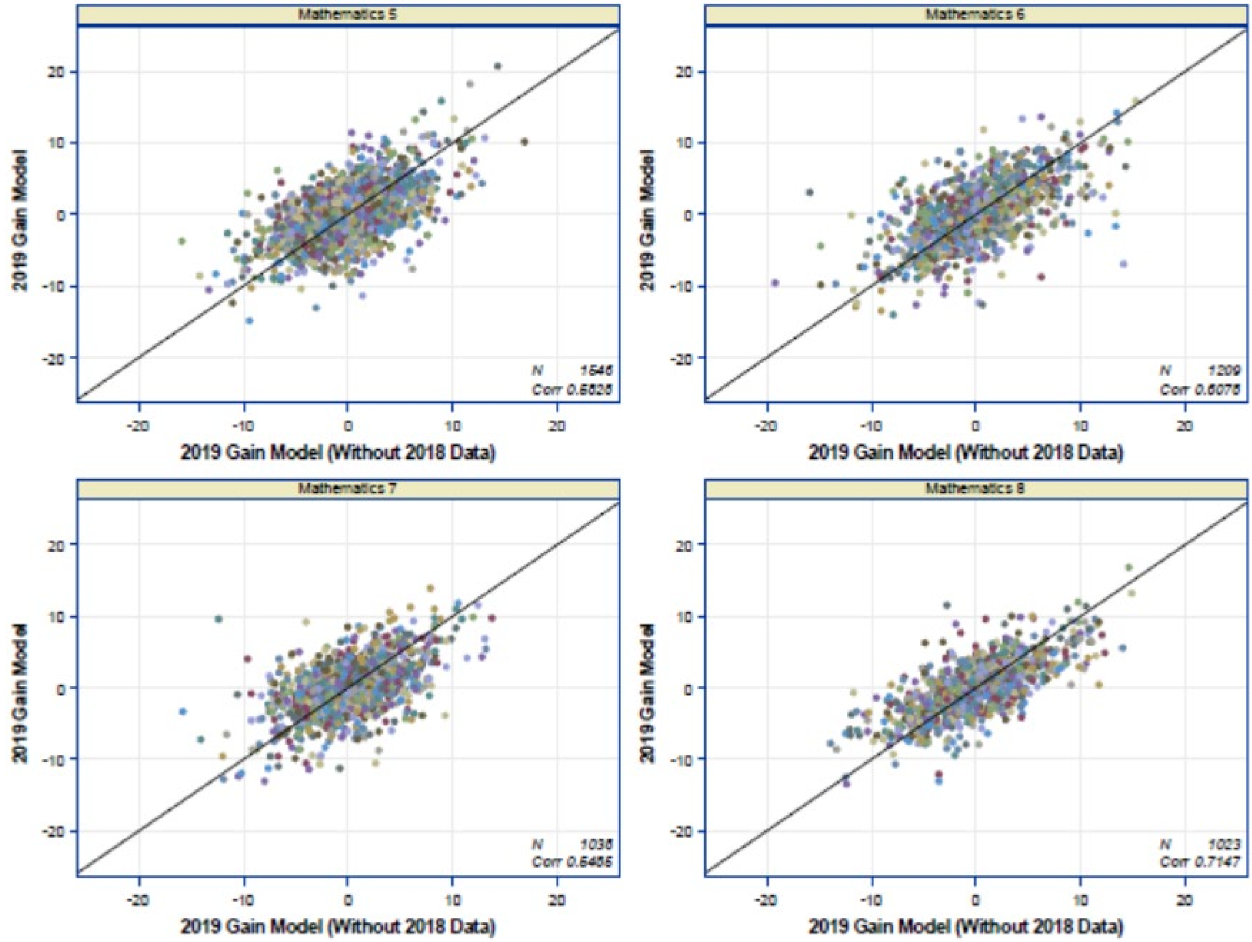


Figure 5: 2019 Gain Model versus 2019 Gain Model without 2018 Data for Teacher-Level Growth Indices in Individual Grades for ELA

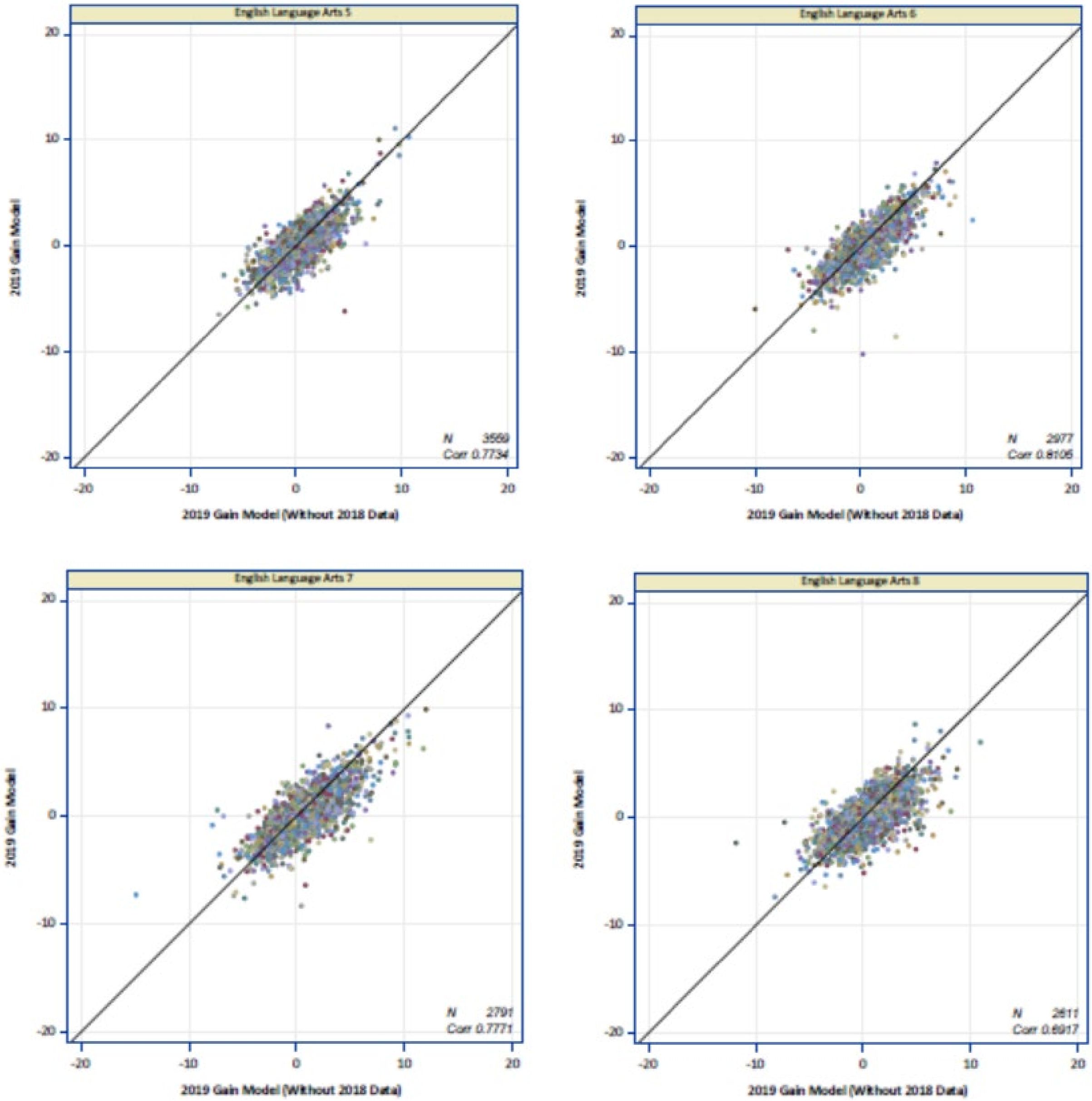
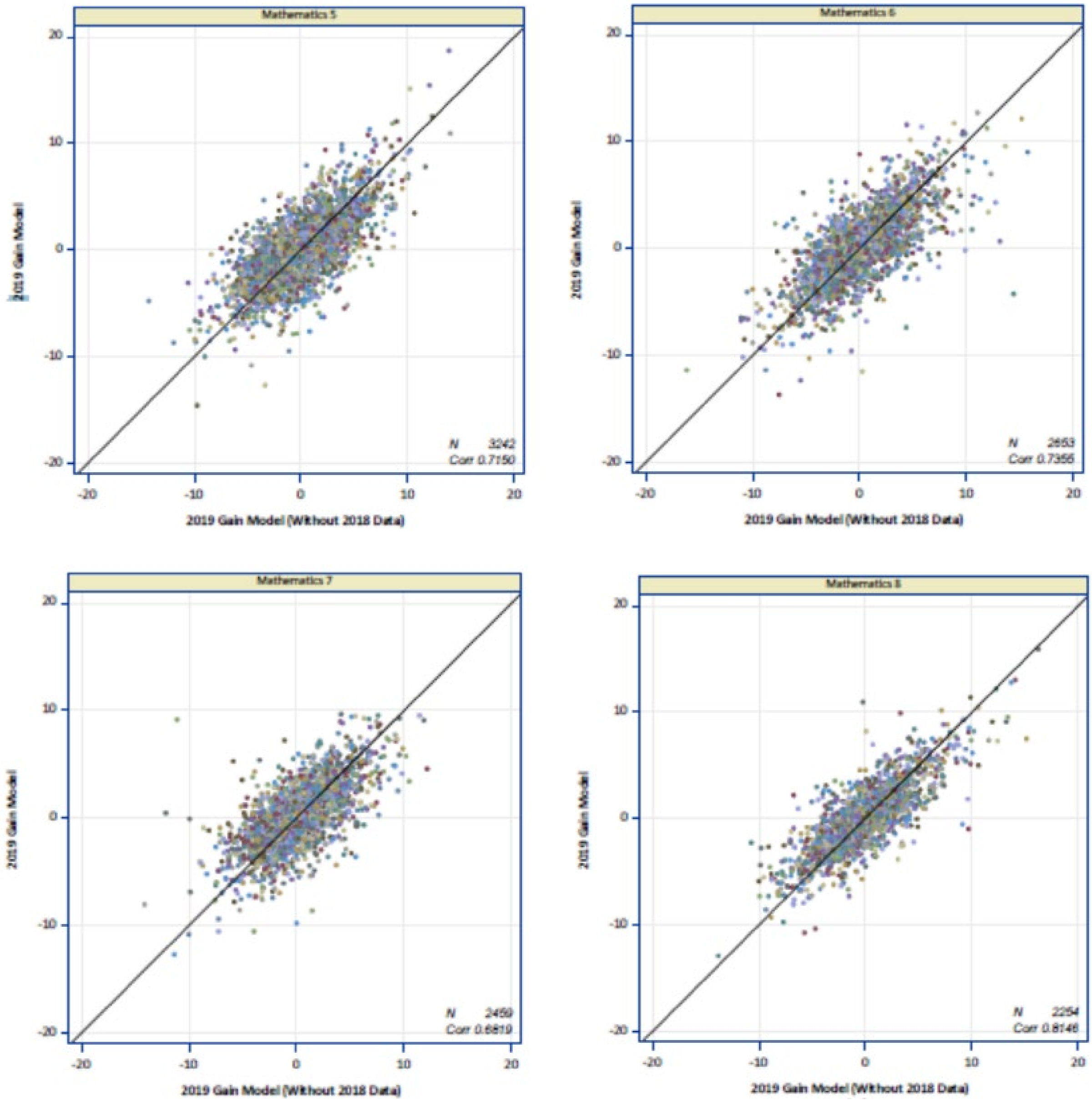


Figure 6: 2019 Gain Model versus 2019 Gain Model without 2018 Data for Teacher-Level Growth Indices in Individual Grades for Math



2.3 Interpreting One- and Two-Year Gains

In moving from a one-year gain to a two-year gain, the change in interpretation relates not only to the time period but also the cohort of students included in the analysis. To explore this idea further, this section provides a simplified example of how gains are calculated with and without a missing year of data in the gain model.

Example 1: Moving from a One-Year Gain to a Two-Year Gain

Table 2 provides the average achievement level for the students testing at a sample district. As a cohort of students moves from one grade to the next, their achievement level can be tracked along a diagonal line. For example, Table 2 shows that the achievement level of Grade 5 students in Year 2 is 25 NCEs and then changes to 36 NCEs when this cohort of students is in Grade 6 in Year 3.

Table 2: Average Achievement in NCEs by Grade and Year for Sample District

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Year 1	13	14	15	16	17	18
Year 2	23	24	25	26	27	28
Year 3	33	34	35	36	37	38

In the computationally ideal situation where all students are present in all three years and students never change districts, the calculation of gains is straightforward. To calculate the gain for Grade 6 in Year 3, it would be the achievement level for Grade 6 in Year 3 minus the achievement level for Grade 5 in Year 2. That would be 36 NCEs minus 25 NCEs, or 11 NCEs.

In the non-computationally ideal situation described above, the gain model calculates means by accounting for missing student scores and allowing for students who move between schools. (More details are available in the Statistical Models and Business Rules document.)

Table 3 provides data for the same district except that data for Year 2 is missing. If there is no Year 2 data, it is not possible to calculate a *one-year gain* for Grade 6 in Year 3. It is possible, however, to calculate a *two-year gain* based on the change in achievement from Grade 4 in Year 1 to Grade 6 in Year 3. This would be 36 NCEs minus 14 NCEs, or 22 NCEs.

Table 3: Average Achievement in NCEs by Grade and Year for Sample District

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Year 1	13	14	15	16	17	18
Year 2	Missing	Missing	Missing	Missing	Missing	Missing
Year 3	33	34	35	36	37	38

For the purposes of this analysis, the two-year gain could then be compared to the sum of the two single-year gains based on a model with Year 2 data. If the missing Grade 5 achievement in Year 2 was 25 NCEs as in Table 2, then this would be (36 NCEs – 25 NCEs) + (25 NCEs – 14 NCEs), which would be 11 NCEs + 11 NCEs, or 22 NCEs. This is the same growth measure as the two-year gain (36 NCEs – 14 NCEs = 22 NCEs). The ideal case is that the two-year gain and the sum of the two single-year gains are the same, and the next model comparison in Section 3 will explore this concept further.

Example 2: Moving from a One-Year Gain to a Two-Year Gain While Changing Schools

Table 4 illustrates the available data in two sample schools: School A serving grades 3–5 and School B serving grades 6–8.

Table 4: Average Achievement in NCEs by Grade and Year for Sample Schools

	School A			School B		
	Grade 3	**Grade 4**	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**
Year 1	13	14	15	16	17	18
Year 2	23	24	25	26	27	28
Year 3	33	34	35	36	37	38

The entering achievement level reported for Grade 6 in Year 3 is an average based on the Grade 6 students' prior achievement from Grade 5 in Year 2. This is relevant for the lowest grade in a school, often Grade 6, since there is no mean *at that school* for the previous grade and year.

In either instance (the computationally ideal situation described in the first example or the average based on students' prior achievement from their prior year schools), there is data available to calculate single-year gains.

However, in Table 5 below, data for Year 2 is missing. For each grade level, the average achievement in NCEs is known for Year 1 and Year 3. Because there is no Year 2 data, it is not possible to calculate a *one-year gain* for Grade 6 in Year 3. It is possible, however, to calculate a *two-year gain* based on the change in achievement from Grade 4 in Year 1 to Grade 6 in Year 3. This would be 36 NCEs minus 14 NCEs, or 22 NCEs.

Table 5: Average Achievement in NCEs by Grade and Year for Sample Schools with Missing Data in Year 2

	School A			School B		
	Grade 3	**Grade 4**	**Grade 5**	**Grade 6**	**Grade 7**	**Grade 8**
Year 1	13	14	15	16	17	18
Year 2	Missing	Missing	Missing	Missing	Missing	Missing
Year 3	33	34	35	36	37	38

From an interpretation perspective, School B's growth measure for Grade 6 in Year 3 includes students' growth from Grade 5 in Year 2 at School A. Although the two-year growth measure is very similar to the two one-year gains, the interpretation has changed.

There are three takeaways from these examples:

- First, the gain model with a missing year of data essentially provides two-year growth measures. As a result, it is reasonable to expect that growth measures with a missing year of data would more closely align with the actual results observed across two years. The following section makes this comparison.
- Second, because the models with a missing year of data (such as the Gain Model without 2018 Data) provide two-year growth measures, the growth measure for grades where students transition from one school to another will then include growth from the feeder school(s) as well as the receiver school. As shown above, in these models, a middle school with grades 6–8 could receive a growth measure for sixth grade based on the students' growth in sixth grade as well as their growth from the feeder elementary school(s) in fifth grade.

  In other words, it is not possible to parse out the individual contribution of the middle school in sixth grade apart from those from the elementary school(s) in fifth grade because of the missing year of test scores. For the district-level growth measures and for the non-transition grades, the two-year growth measures are still solely representative of growth within the specific district and the non-transition grades for the school are still solely representative of growth within the specific school.
- The third takeaway is that, at a particular school, the growth of certain groups of students are not represented in the two-year measures as they would be in two one-year growth measures. In the example above, it is not possible to measure the growth of Grade 4 students in Year 3 at School A. Because there is no Grade 3 data in Year 2 (and no statewide assessment from Grades K–2 available), it is not possible to report Grade 4 growth from Year 3 or include in any composite calculations. Similarly, it is not possible to report Grade 8 growth from Year 2 or

include in any composite calculations because there is no exiting achievement for these students in their last year at the school.

To say this differently, in comparing results across grades from a two-year measure with missing data and from two one-year growth measures without missing data, the gains will likely align well in most subjects and grades. However, they will not align exactly because there will be two cohorts of students that are missing from the latter model: the Grade 3 students in the year of missing data and the Grade 8 students in the year of missing data.

For these reasons, the interpretation of one-year and two-year gains varies, and there is some nuance in what the growth measure represents. The next two sections explore two-year models, both with and without a missing year of data.

3 2019 Gain Model (Two-Year) versus 2019 Gain Model without 2018 Data

3.1 Overview

As stated in the previous section, a key difference between the two models is that one model was based on a one-year gain while the other was based on a two-year gain. To provide a more comparable analysis, this section compares a sum of single-year 2017-18 and 2018-19 growth measures based on the gain model *with* 2017-18 test scores to the 2018-19 growth measures based on the gain model *without* 2017-18 test scores. The former model could be considered the standard or typical model for a two-year gain of EVAAS reporting in the consecutive grade-given tests. The latter also uses the standard model but the immediate prior year of test scores is missing, so the growth measures represent a two-year gain. The point of this comparison is to provide context to how the sum of two single-year gains compares to a two-year gain with missing data for the middle year.

Table 4: Overview of Two Models in Comparison

Comparison	2019 Gain Model (Two-Year)	2019 Gain Model without 2018 Data
Model approach	Gain model	Gain model
One-year growth measure	No	No
Gains measure growth from	Sum of 2016-17 to 2017-18 and 2017-18 to 2018-19	2016-17 to 2018-19
Inclusion of 2017-18 test scores	Yes	No

For this comparison, correlations and scatterplots can provide insight into the relationship between the two models.

It is clear from the correlations and scatterplots that the results from these two models are very similar, with a correlation above 0.99 across all district and school results for ELA and Math for all individual subject/grades. In comparing results from this section to the previous one, the findings make sense as they both represent a growth measure from a two-year period.

Note that, for this comparison, only district and school models were analyzed as it is highly unusual for teachers to teach the same students two years in a row although it is fairly common for districts and schools to have instructional responsibility for many of the same students two years in a row.

Correlations for individual subjects and grades are available in the following scatterplots, which plot the district and school growth index for individual subject/grades for ELA and Mathematics. Results for 2019 Gain Model (Two-Year) are on the Y axis, and results for 2019 Gain Model without 2018 Data are on the X axis. The different colors in the graph do not have an assigned meaning; they are simply there to help interpret the distribution more easily.

3.2 District and School Results

Figure 7: 2019 Gain Model (Two-Year) versus 2019 Gain Model without 2018 Data for District-Level Growth Indices by Subject/Grade for ELA

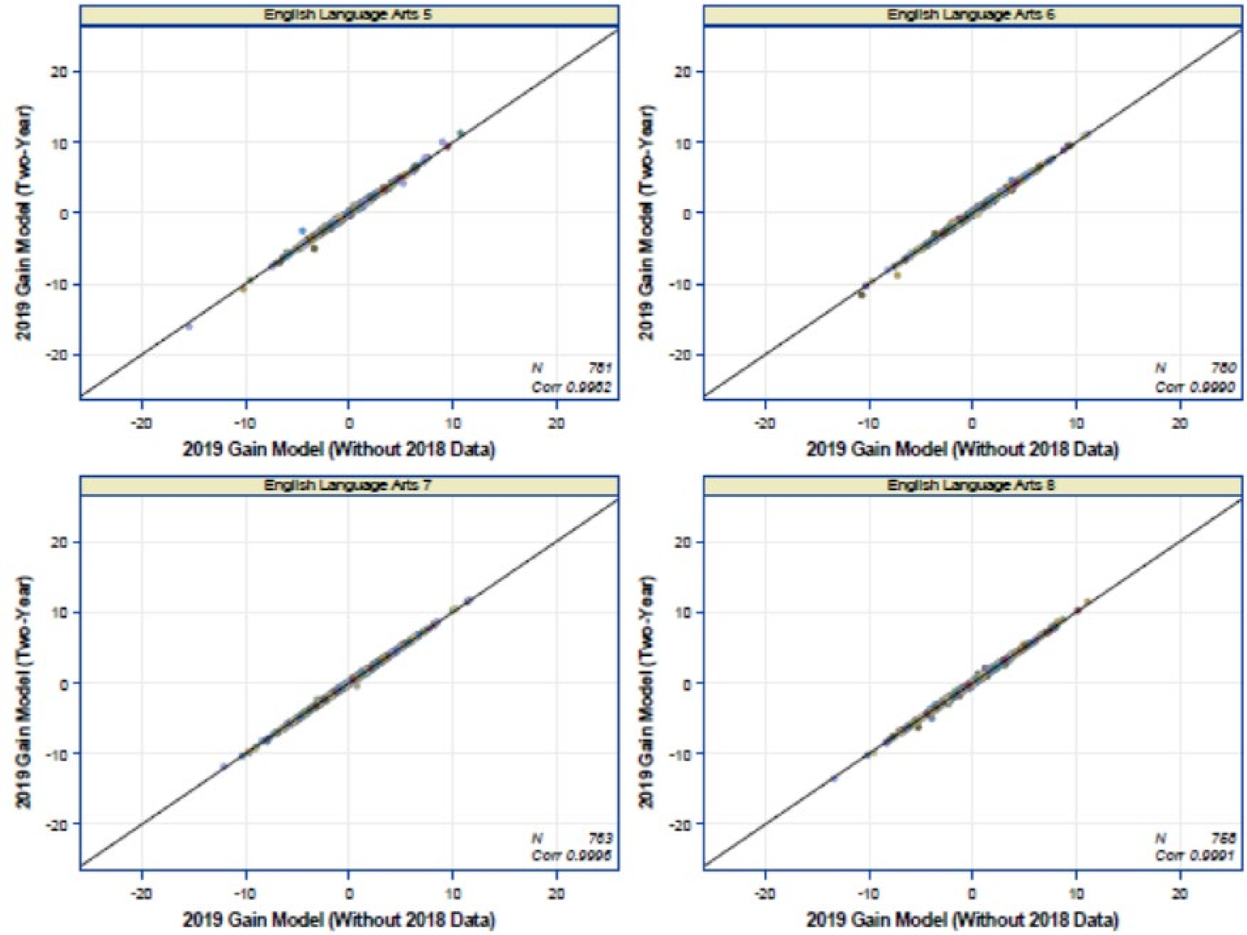


Figure 8: 2019 Gain Model (Two-Year) versus 2019 Gain Model without 2018 Data for District-Level Growth Indices by Subject/Grade for Math

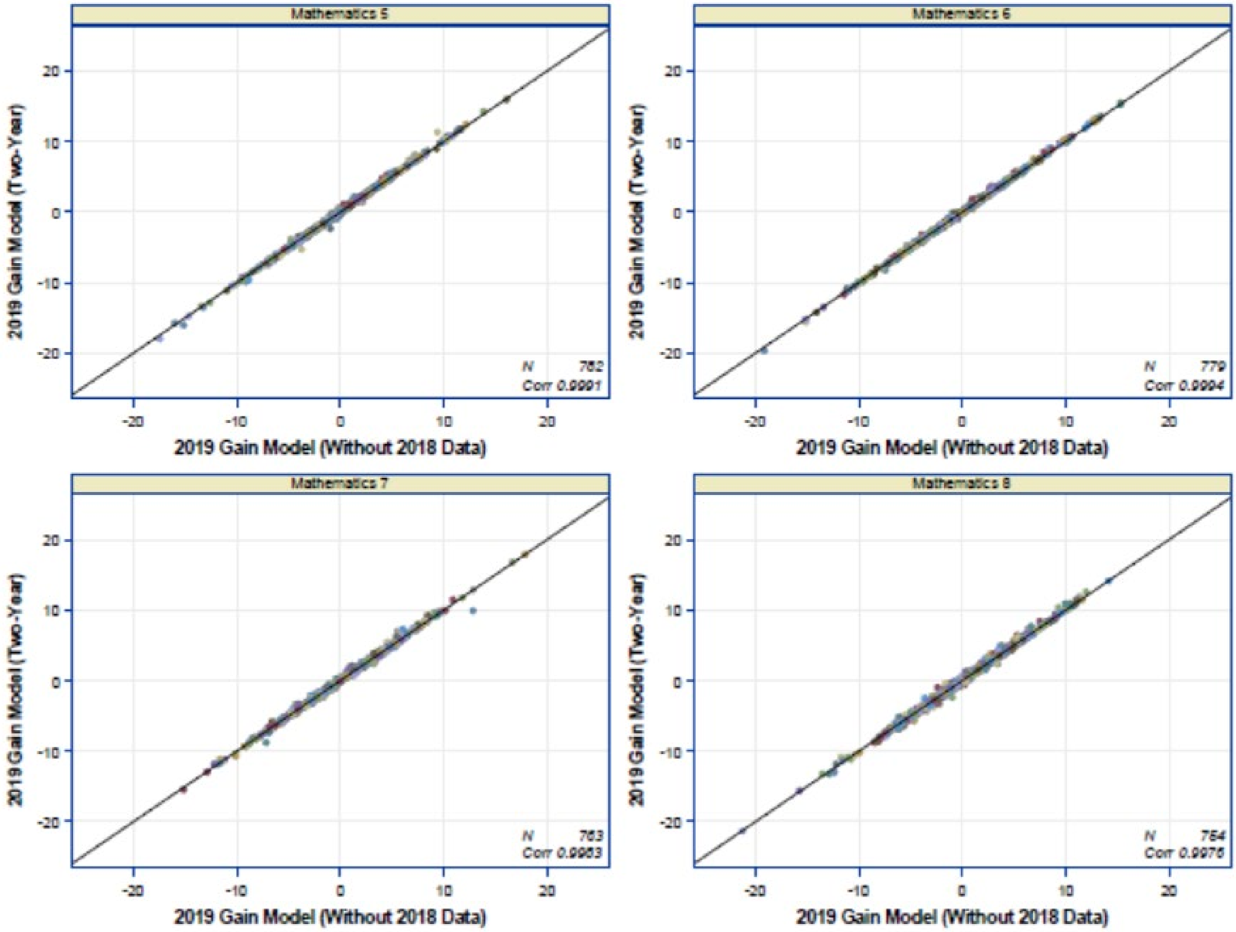


Figure 9: 2019 Gain Model (Two-Year) versus 2019 Gain Model without 2018 Data for School-Level Growth Indices by Subject/Grade for ELA

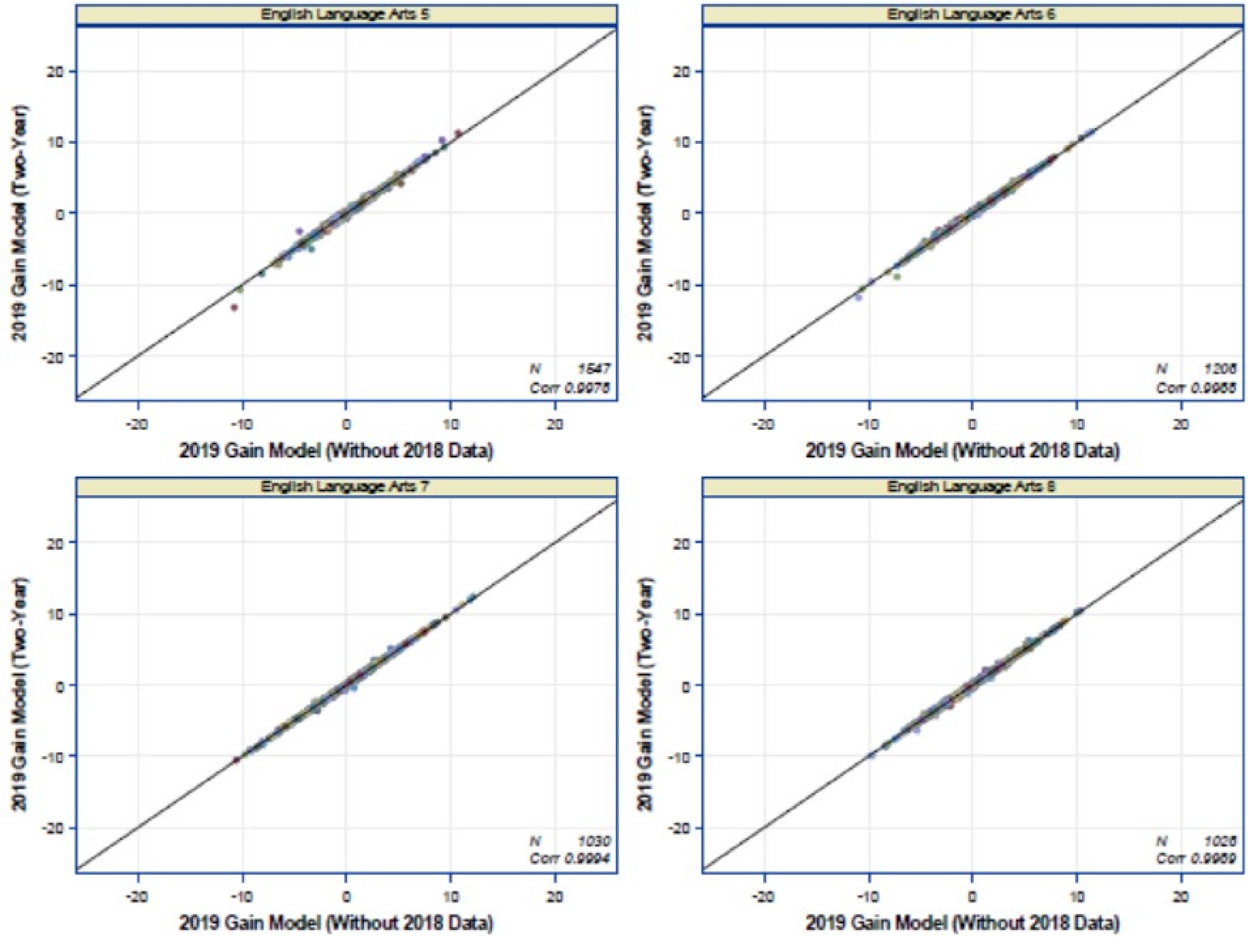
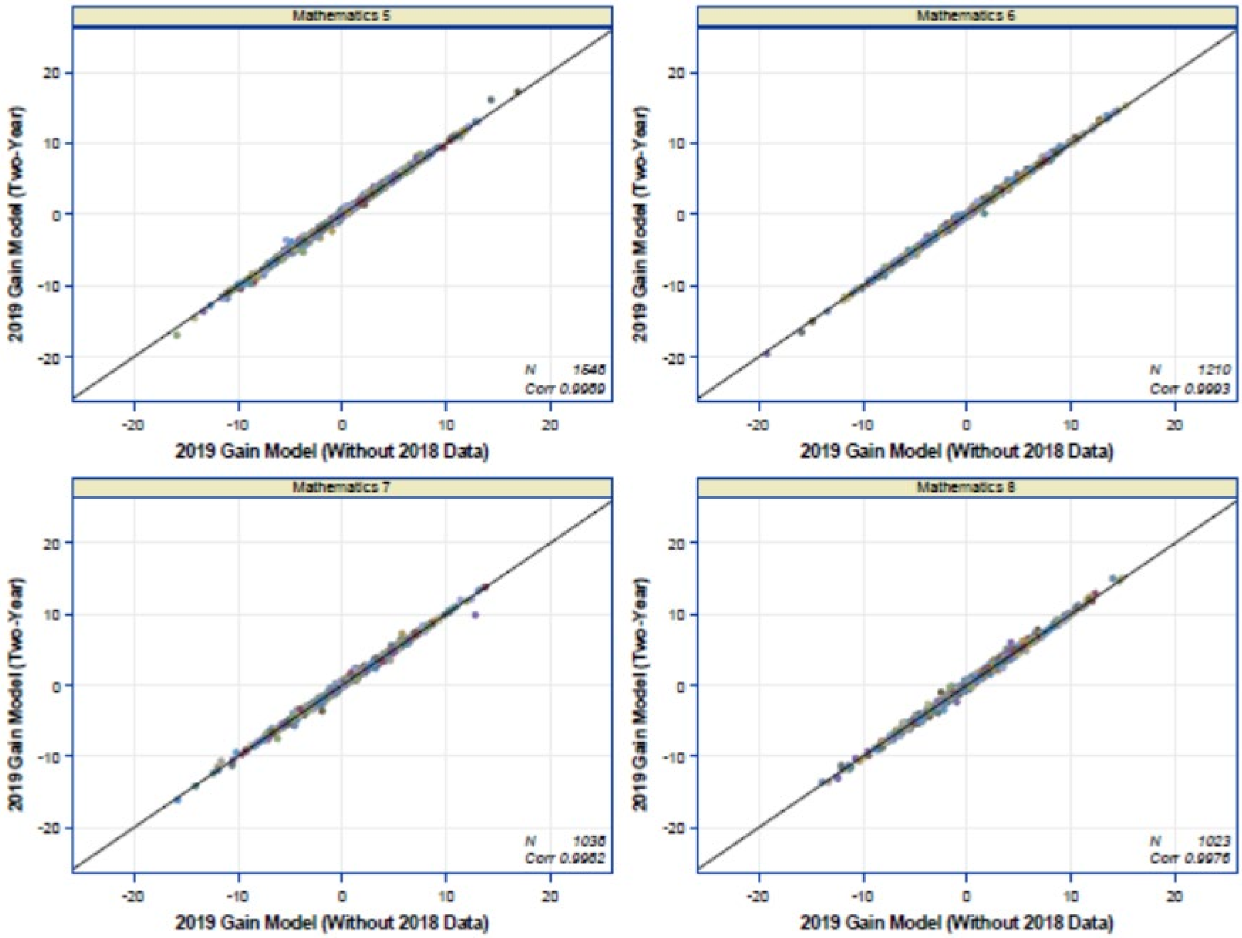


Figure 10: 2019 Gain Model (Two-Year) versus 2019 Gain Model without 2018 Data for School-Level Growth Indices by Subject/Grade for Math



4 2019 Predictive Model versus 2019 Predictive Model without 2018 Data

4.1 Overview

This section compares 2018-19 predictive model growth measures *with* 2017-18 test scores to 2018-19 predictive model growth measures *without* 2017-18 test scores. The point of this comparison is to provide context to how removing the immediate prior year of test scores can impact the predictive model. Note that the predictive model is used to measure growth for assessments given in non-consecutive grades, such as OST Science in grade 8 as well as the high school end-of-course assessments in Algebra I, English Language Arts, Geometry, and Mathematics I and II. Because these assessments are not administered every year, it is possible that students do not have any test scores in the *immediate* prior year. The model can provide a robust estimate of students' entering achievement for the course by using all other available test scores from other subjects, grades, and years.

In other words, the predictive model in this comparison did not require any technical adaptations to account for the missing year of data; any predictors from the previous year were just excluded from the model. As mentioned above, this comparison is useful for understanding how an entire year of missing data could impact the 2020-21 growth measures.

Table 5: Overview of Two Models in Comparison

Comparison	2019 Predictive Model	2019 Predictive Model without 2018 Data
Model approach	Predictive Model	Predictive Model
One-year growth measure	Yes	No
Gains measure growth from	2017-18 to 2018-19	2016-17 to 2018-19
Inclusion of 2017-18 test scores	Yes	No

For this comparison, correlations and scatterplots can provide insight into the relationship between the two models.

The district, school, and teacher growth indices in all courses demonstrate a very strong correlation between the 2019 Predictive Model and 2019 Predictive Model without 2018 Data.

- The district correlations range from 0.92 to 0.98.
- The school correlations range from 0.93 to 0.97.
- The teacher correlations range from 0.93 to 0.98.

These are almost as high the correlations from the second comparison, 2019 Gain Model (Two-Year) and 2019 Gain Model without 2018 Data, and they are very close to +1. In contrast to the gain model, the predictive model is typically used for assessments that are not administered in consecutive grade levels. As a result, although fewer prior scores are available to be used as predictors with a missing year of data, the interpretation is more similar to a typical year than for the gain model.

Correlations for individual courses are available in the following scatterplots, which plot the district, school, and teacher growth index for individual courses. Results for 2019 Predictive Model with 2017-18 test scores are on the Y axis, and results for 2019 Predictive Model without 2018 Data are on the X axis. The different colors in the graph do not have an assigned meaning; they are simply there to help interpret distribution more easily.

4.2 District, School and Teacher Results

Figure 11: 2019 Predictive Model versus 2019 Predictive Model without 2018 Data for District-Level Growth Indices by Course

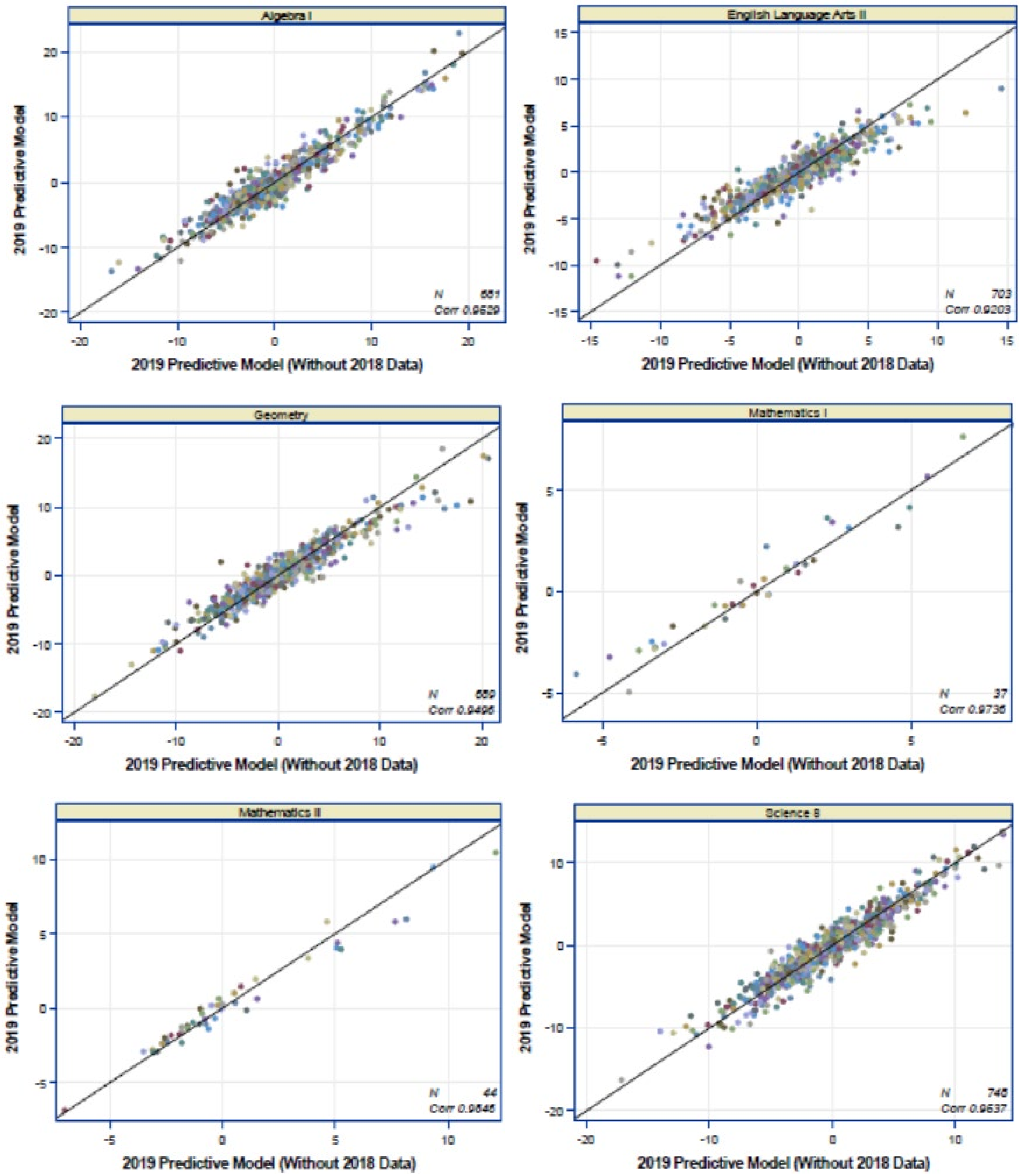


Figure 12: 2019 Predictive Model versus 2019 Predictive Model without 2018 Data for School-Level Growth Indices by Course

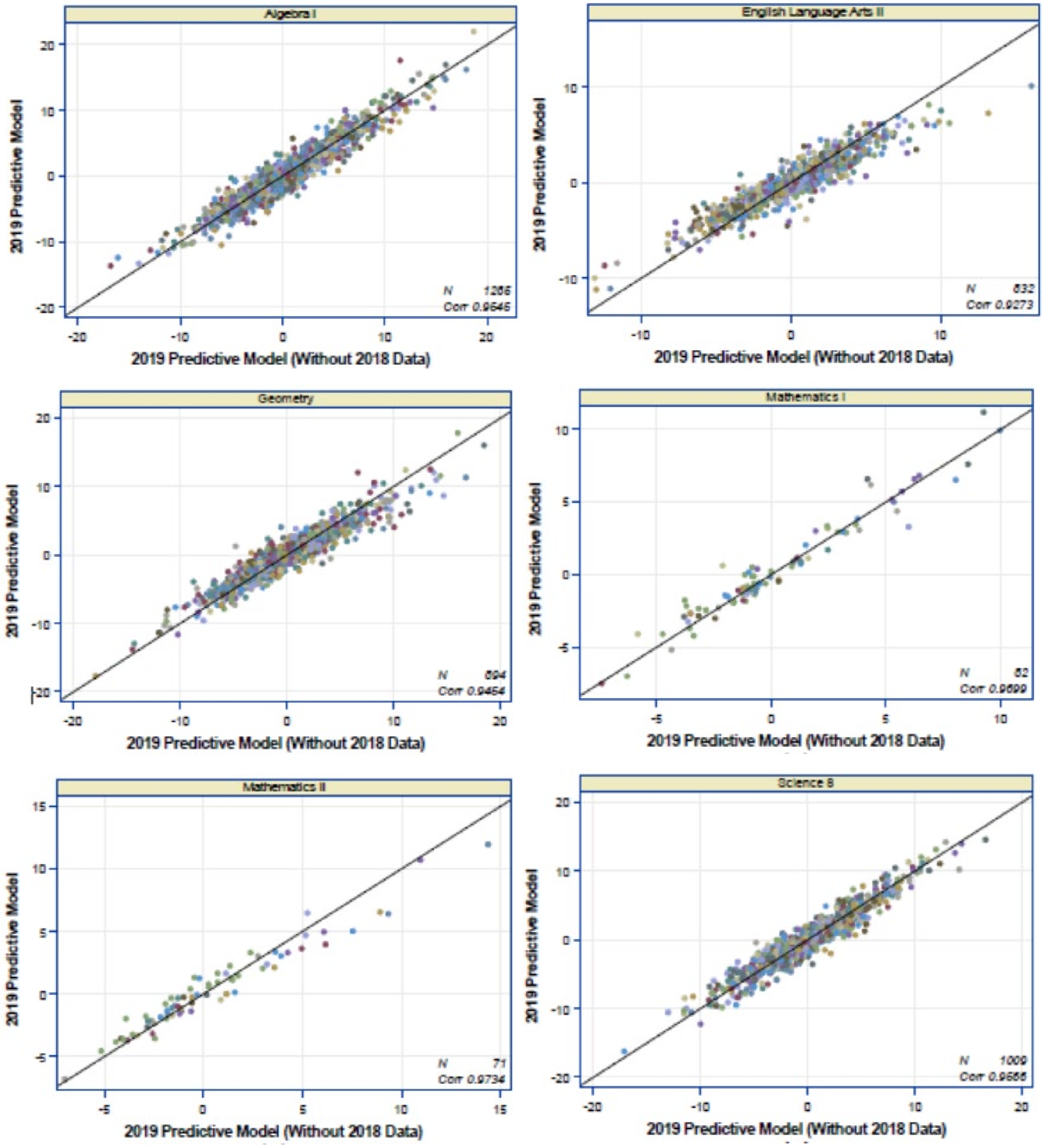
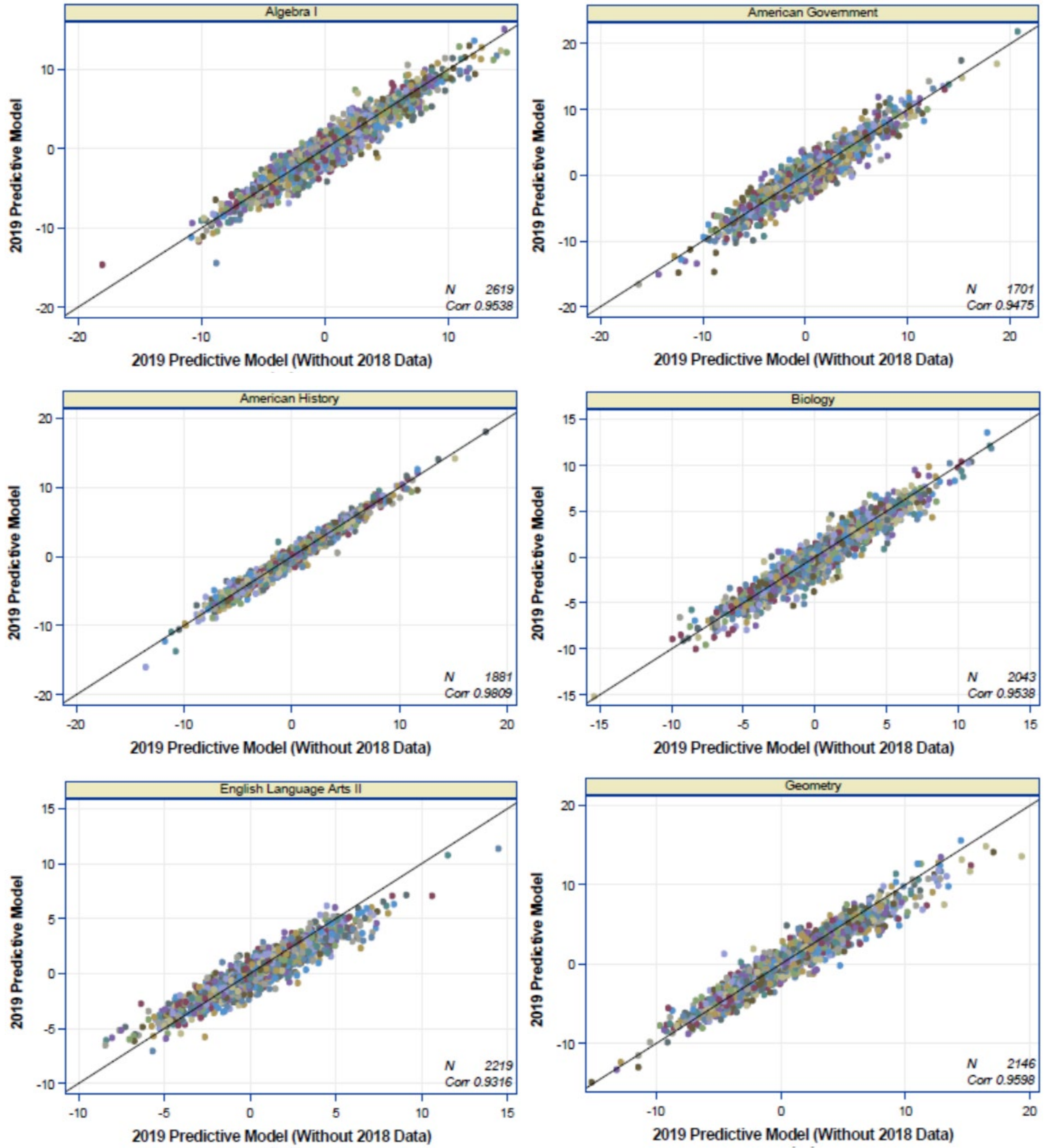
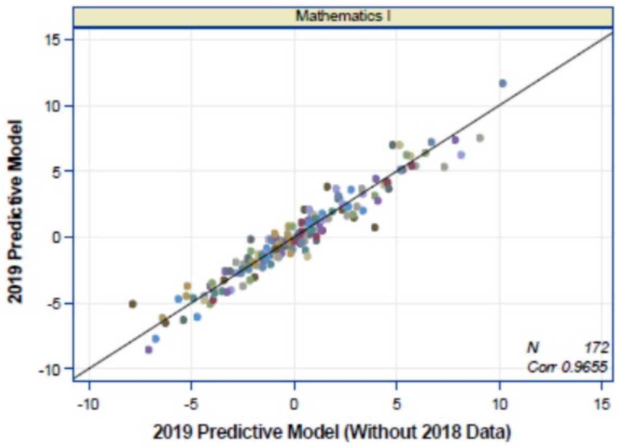
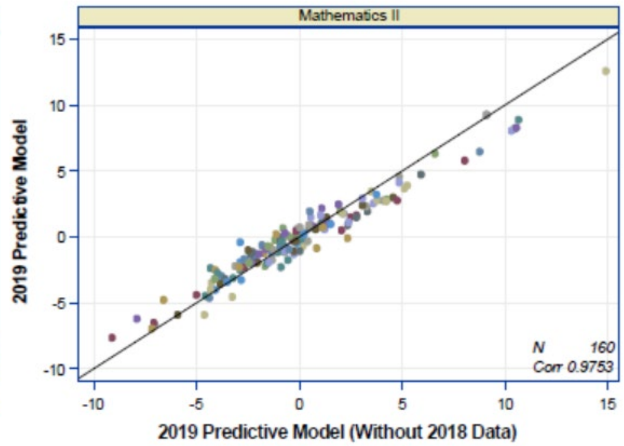
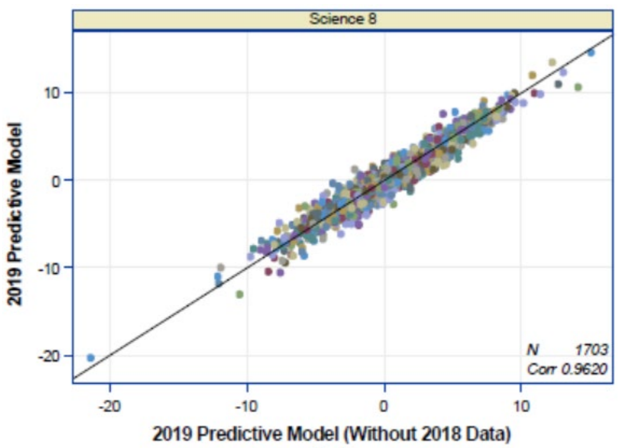


Figure 13: 2019 Predictive Model versus 2019 Predictive Model without 2018 Data for Teacher-Level Growth Indices by Course




Mathematics I
2019 Predictive Model
2019 Predictive Model (Without 2018 Data)
N 172
Corr 0.9655


Mathematics II
2019 Predictive Model
2019 Predictive Model (Without 2018 Data)
N 160
Corr 0.9753


Science 8
2019 Predictive Model
2019 Predictive Model (Without 2018 Data)
N 1703
Corr 0.9620

5 Conclusions and Additional Considerations

The comparisons of models without a missing year of data to models with a missing year of data illustrated that results for all of the comparisons were highly correlated. More specifically:

- A comparison between the standard, one-year gain model used for consecutive grade-given tests and gain model results with a missing prior year of data illustrates positive correlations for all individual grades and subjects for districts (0.48 to 0.71), schools (0.49 to 0.71), and teachers (0.68 to 0.81). In this case, one of the comparisons is essentially based on a two-year gain, whereas the other is based on a one-year gain.
- To provide a more comparable analysis, the analysis also compared a sum of single-year 2017-18 and 2018-19 growth measures based on the gain model *with* 2017-18 test scores to the 2018-19 growth measures based on the gain model *without* 2017-18 test scores. The results from these two models are very similar, with a correlation above 0.99 across all district and school results for ELA and Math for all individual subject/grades. This finding is not surprising as both models represent a growth measure from a two-year period.
- The predictive model is used to measure growth for assessments given in non-consecutive grades, such as OST Science in grade 8 as well as the high school end-of-course assessments in Algebra I, English Language Arts, Geometry, and Mathematics I and II. The analysis for these subjects compared predictive model results to predictive model results that did not consider the prior year of testing data. The district (0.92 to 0.98), school (0.93 to 0.97), and teacher (0.93 to 0.98) correlations show that the actual results are highly correlated with results with a missing year of data.

The analysis raises considerations for the interpretation of district, school, and teacher results from the gain model used for consecutive grade-given tests when there is a missing prior year of data:

- With a missing prior year of data, results from the gain model can be considered to measure growth observed over a two-year period. To the extent that districts, schools, and teachers served students in the current year but not the prior year, the interpretation of the growth measure can include students' experiences with a different district, school, or teacher in the previous year.
- In contrast to the gain model, the predictive model is typically used for assessments that are not administered in consecutive grade levels. As a result, although fewer prior scores are available to be used as predictors with a missing year of data, the interpretation is more similar to a typical year than for the gain model.

As a final reminder, the simulations included in this report are focused solely on the missing year of data and do not estimate the pandemic's impact on student learning in districts, schools, and classrooms. Although the results will enhance ODE's understanding of how growth models might be impacted by a missing year of assessment data, more analysis is possible when the 2020-21 assessment data becomes available, which might bring forth additional relevant considerations.

In terms of what these findings mean for educators and administrators, the key takeaway is that 2020-21 growth measures are comparable to a two-year measure for the gain model and to a one-year measure for the predictive model. While the interpretation varies in 2020-21 reporting compared to previous years, these measures are still a useful resource for educators, administrators, policymakers and other stakeholders. They can provide insights and answer questions such as the following:

- How did we help our students grow in comparison to students in other schools and districts across the state during these unusual times?

- Did some student populations have more success than others? How did we contribute to that success, and how can we apply what we learned and apply it to future instructional programming?
- Did some student populations have less success? What factors contributed to those results, and how can we help students recover from any incomplete learning?

This information can help educators target their resources more efficiently, identify exemplars, and make data-driven decisions for future years. SAS will work with ODE to provide additional resources for educators and other stakeholders based on the 2020-21 reporting to assist with interpretation and context.

6 Appendix

6.1 Gain Model

Historically, EVAAS growth reporting has used a gain-based model for the tests given in consecutive grades, such as OST Mathematics and Reading for grades 3–8. Known as the gain model or more formally as the multivariate response model (MRM), it can also be described as a linear mixed model or repeated measures model. The gain model measures growth between two points in time for a group of students, and the growth expectation is met when a cohort of students from grade to grade maintains the same relative position with respect to statewide student achievement for a specific subject and grade.

In the gain model, the growth measures themselves are not based on simple gains or a simple average of differences in student achievement over time. A more robust approach is required to address non-trivial complexities associated with testing data, like the fact that missing test scores are not random or that all test scores have measurement error. Particularly relevant to this investigation is that the gain model includes all students with valid data even if they have missing test scores and includes multiple years of testing history across multiple subjects for each individual student. This means that EVAAS growth measures for 2018-19 grade 6 Math incorporated available test scores from Mathematics and Reading for the students as 2018-19 sixth graders and 2017-18 fifth graders as well as 2020-21 fourth graders and 2015-16 third graders.

EVAAS reporting is based on three separate analyses for the gain model: one each for districts, schools, and teachers. The district and school models are essentially the same; they perform well with the large numbers of students that are characteristic of districts and most schools. The teacher model uses a different approach that is more appropriate with the smaller numbers of students typically found in teachers' classrooms. The district and school models use students flagged as accountable, and the teacher model uses students who tested and are linked to the teacher.

Without 2019-20 test scores, it was impossible for the gain model to determine whether students maintained the same relative position in 2019-20 compared to 2018-19 since there was no current year achievement metric available to use. The analyses in this document explore the feasibility of using prior years' data (test scores from 2018-19 and prior years) to assess whether students maintained the same relative position in 2020-21 compared to 2018-19 with the gain model approach.

6.2 Predictive Model

Tests that are not administered to students in consecutive years require a different modeling approach from the gain model. Such tests include OST Science in grades 5 and 8 and end-of-course tests and college readiness tests like the ACT. This modeling approach is called the predictive model or univariate response model (URM). It is also a linear mixed model and can be further described as an analysis of covariance (ANCOVA) model. The predictive model measures the difference between students' predicted scores for a particular subject/year with their observed scores. The growth expectation is met when, on average, the observed scores for a group of students with a district/school/teacher met their predicted scores.

An important distinction between the gain model and the predictive model is that the gain model includes all students in its models while the predictive model includes only students who have at least three prior test scores. These prior test scores can be from any subject and grade, and the model will include more prior test scores if available. However, a student needs at least three prior test scores in order to receive a predicted score for the particular subject/year. This means that, like the gain model, the predictive model can accommodate students with missing test scores and does not require that students have the same set of prior test scores to be included in the model.

EVAAS reporting in the non-consecutive grade-given tests is also based on three separate analyses based on the predictive model, one each for districts, schools, and teachers. The three models are essentially the same. Note that these predictive models, like the teacher gain model, use shrinkage estimation ("random effects"). This is true for all three predictive models (district/school/teacher) though the amount of shrinkage is typically small in the case of the school model and, especially, the district model. The district and school models use students flagged as accountable, and the teacher model uses students who tested and are linked to the teacher.

6.3 Growth Expectation for Both Models

Although the gain and predictive models have different ways to measure growth, the growth expectation in these models is conceptually similar. In both models, the growth expectation represents the average amount of progress observed in the state in the current year.

Typically, the growth expectation is set at zero, such that *positive* gains or effects are evidence that students made *more* than the expected growth and *negative* gains or effects are evidence that students made *less* than the expected growth.

More precisely, the gain model defines the growth expectation as students maintaining the same relative position with respect to the statewide student achievement from one year to the next in the same subject area. For example, if students' achievement was at the 50^{th} NCE in 2018 grade 4 Math, based on the 2018 grade 4 Math statewide distribution of student achievement, and their achievement is at the 50^{th} NCE in 2019 grade 5 Math, based on the 2019 grade 5 Math statewide distribution of student achievement, then their estimated gain is 0.0 NCEs. This means that students, on average, met the growth expectation.

The predictive model defines the growth expectation as students with a district, school, or teacher making the same amount of progress as students with the average district, school, or teacher in the state for that same year/subject/grade. For example, if students' predicted scores were 715 for Algebra I based on their prior test scores and their observed scores were also 715 for Algebra I, then the difference is 0 scale score points. This means that students, on average, met the growth expectation.