URM Modeling Approach for Value-Added

Introduction

This document provides an overview of the value-added modeling approach used for end-of-course (EOC), Science, and Social Studies tests. Tests that are not given for consecutive years require a different modeling approach from the gain-based approach (known as the MRM) used for OST Math and Reading. To be more specific, although some EOC tests are taken in succession, not all students take them in the same grades, and this approach is more appropriate in those scenarios.

For Ohio’s value-added reporting, the modeling approach for EOC, Science and Social Studies is called the univariate response model (URM). The URM is a regression-based model, which measures the difference between students’ predicted scores for a subject/year with their observed scores. Students’ predicted scores are based on the relationships between the test in question and prior test scores, meaning there must be a sufficient relationship between these scores to provide a predicted score. The growth expectation is met when, on average, students linked to a specific district/school/teacher made the same amount of growth as students linked to the average district/school/teacher with the state for that same year/subject/grade.

Key Technical Aspects of the URM Approach

The URM approach shares many of the statistical advantages as the MRM approach. The URM uses all available testing data for each student, accommodates teaching scenarios where more than one teacher has responsibility for a student’s learning in a specific subject/grade/year, and does not require all test scores to be on the same scale. Although some value-added/growth models require students to have all predictors or the same set of predictors, the URM does not. Students can be included if they have at least three prior test scores in any subject/grade. As mentioned above, the URM does not require consecutive grade given tests while the MRM does.

Common Questions for Interpretation

What is student growth? Student growth looks at how much academic growth students make within a tested subject/grade over time.

What is the modeling approach? For tests given in non-consecutive grades, growth is measured through a regression-based approach, which compares students’ predicted scores with their observed scores. A student’s predicted score is obtained by looking at his or her prior testing performance (scores) across all subjects, as well as all students with similar prior testing performance, and estimating how that group of students scored, on average, on that test.

How is a student’s predicted score determined? The process of generating each student’s predicted score begins with building a robust statistical model of all students who took the selected assessment in the most recent year. The model includes the scores of all students in the state or reference group, along with their testing histories across years, grades, and subjects. This model determines the relationships of all the prior assessment scores to students’ performance on the selected assessment.
This model determines the relationships between all students’ prior assessment scores and their actual performances on the selected assessment. By considering how all other students performed on the assessment in relation to their testing histories, the model is able to calculate a predicted score for each student based on his or her individual testing history.

The model uses prior test scores from all grades and subjects as this improves the reliability in determining the expected score. (For example, grade 8 Math, Reading, and Science scores are typically related to the Algebra I scores.) Furthermore, the prior test scores are weighted according to the strength of the relationship that they have with the end-of-course subject, so Math scores can be weighted more than Reading scores for the Algebra I prediction although reading scores can still help improve the prediction.

Conceptually, it might be useful to consider the expectation as the student’s entering achievement since it is based on all the student’s achievement data prior to the year in question.

**What is the growth measure?** The growth measure is related to the number of scale score points a group of students scored above or below their predicted score, which considers their prior testing performance. The growth measure is not an exact difference between the predicted and observed scores for a group of students. To provide additional protections to educators, the growth measure uses shrinkage estimation, which means that each district/school/teacher is assumed to have the average growth measure (zero) until the weight of the evidence pulls the growth measure above or below zero. In other words, the growth measure is a function of the difference between the predicted and observed scores for a group of students. This function will consider the number of students used in the growth measure as well as the variability of the districts/schools/teachers across that entire subject and grade for a given year. The more students a teacher has, the more certainty there is on students’ growth, and this means less shrinkage. When less variability among school or teacher growth measures is observed throughout the state, more shrinkage will be applied to those school or teacher growth measures. Finally, at the teacher level, this function also accounts for the sharing of instructional responsibility.

**How is the growth measure calculated?** The key steps are as follows:

- Each student’s prior testing performance is used to determine his or her predicted score.
- The predicted score assumes that each student has the average schooling experience for the reference population.
- The predicted score is compared to the actual score for each student.
- The growth measure is a function of the difference of the predicted and observed scores for a group of students linked to a teacher, school, or district.

**What is the standard error of the growth measure?** This is a measure of certainty of the estimated school or teacher growth measure, and it provides insight regarding the certainty that a growth measure is decidedly above or below expected growth protecting educators and schools from misclassification. The standard error is based on both the quantity and quality of data for the school or teacher, such as the number of students linked to the teacher or school and the completeness of the students’ test scores. Note, standard error relates to certainty around a teacher’s individual growth measure and is different for each teacher. Therefore, the standard error is not forcing any preset percentages of teachers into the different categories. In other words, the standard error is not the standard deviation of the teacher value-added measures.

**What is the growth index?** The growth index is then calculated by dividing the growth measure by its standard error in order to determine the significance of the growth measure.
What is the “the average schooling experience”? For School Value-Added reports, the average schooling experience means the growth that students made in the average school from the reference population for that course in the current year. In other words, for Algebra I School Value-Added reports, the average schooling experience represents the progress that Algebra I students made in the average school in the reference population. This reference group is based on the pool of test-takers, which is statewide for state assessments and a smaller subset for vendor assessments. The average schooling experience considers each student’s prior testing performance.

For Teacher Value-Added reports, the average schooling experience means the growth that students made who were linked to the average teacher from the reference population for that course in the current year. This ensures an “apples-to-apples” comparison among schools and teachers.

What is the expected growth and how is it calculated? Expected growth means that the observed scores of a group of students linked to a teacher/school/district were, on average, the same as their predicted scores. In other words, the difference between students’ predicted and observed scores was zero (or close to zero).

It is possible to have a small negative effect or small positive effect yet still be considered “making expected growth.” Because the growth measure is an estimate of student growth, there must be sufficient evidence that a growth estimate is decidedly above or below expected growth to receive the green or red labels.

At a Glance - Understanding the URM Value-Added Report

At a glance, you can see whether students with a teacher, school or district made more, less or about the expected growth due to the color-coding. These colors can be interpreted as follows:

- **DG**: Significant evidence that students made more progress than the Growth Standard.
- **LG**: Moderate evidence that students made more progress than the Growth Standard.
- **Y**: Evidence that students made progress similar to the Growth Standard.
- **O**: Moderate evidence that students made less progress than the Growth Standard.
- **R**: Significant evidence that students made less progress than the Growth Standard.

This information provides educators with a quick indicator of whether they are meeting the academic needs of all their students. Other EVAAS reports (Diagnostic reports, Decision Dashboard, etc.) can provide insight as to why this is the case.

Deeper Dive - Understanding the URM Value-Added Report

The Value-Added reports will be available through the EVAAS web application. The reports are designed to provide basic information about the analysis (number of students included, students’ entering achievement, etc.) as well as to provide quick feedback on students’ growth over time.
Figure 1: Sample URM School Value-Added Report

Report: School Value-Added  Test: Ohio’s State Tests End of Course
School: Coral Snake High School  Subject: English Language Arts II
District: Big City School District  Type: Accountable
Year: 2025

The OST Tested reports include all students who were tested at the school or district. Reporting that includes only those students who were accountable to that school or district is available under the OST Accountable test.

The values in the table below are rounded for display purposes. Note that the unrounded gain is divided by the unrounded standard error to obtain a gain index. The gain index is rounded before being displayed on the Progress component of the Ohio School Report Card. For this reason, the numbers on the two reports might be very different.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Year</th>
<th>Number of Students</th>
<th>Average Score</th>
<th>Average Percentile</th>
<th>Average Predicted Score</th>
<th>Average Predicted Percentile</th>
<th>Growth Measure</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Language Arts II</td>
<td>2013</td>
<td>1566</td>
<td>1541.4</td>
<td>61</td>
<td>1541.5</td>
<td>66</td>
<td>0.35</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>2014</td>
<td>1630</td>
<td>1657.3</td>
<td>64</td>
<td>1681.3</td>
<td>72</td>
<td>-0.9</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>1681</td>
<td>1600.0</td>
<td>51</td>
<td>1465.4</td>
<td>61</td>
<td>3.6 LG</td>
<td>3.6</td>
</tr>
<tr>
<td>Yr Avg</td>
<td></td>
<td>4657</td>
<td>1532.6</td>
<td>59</td>
<td>1513.7</td>
<td>60</td>
<td>-1.0 Y</td>
<td>1.0</td>
</tr>
</tbody>
</table>

- **Number of Students** is the number of students included in this report. This number may differ from the number of students who tested since students must have sufficient historical data to be included in the analysis.

- **Average Score** displays the average observed (actual) scale scores or Normal Curve Equivalents (NCEs) of all students who are included in the analysis for the school.

- **Average Percentile** displays where the Average Score falls in the distribution of all students’ observed (actual) scores in the reference population.

- **Average Predicted Score** displays the average predicted scale scores or Normal Curve Equivalents (NCEs) of all students who are included in the analysis for the school.

- **Average Predicted Percentile** displays where the Average Predicted Score falls in the distribution of all students’ predicted scores in the reference population.

- **Growth Measure** is an estimate of the school effect, which compares student progress to the average school in the reference population.

- **Standard Error** is an expression of certainty. Growth is reported as an estimate, and its significance depends on the amount of certainty in the estimate.
Figure 2: Sample URM Teacher Value-Added Report

- **Growth Measure** is a conservative estimate of the academic progress of the students linked to the teacher.
- **Standard Error** is a measure of certainty around the Growth Measure.
- **Index** is the growth measure divided by the standard error.
- **Level** is determined by the index based on the rules shown below in the legend.
- **Number of Students** shows the number of students included in the analyses. This number might vary from the number of students the teacher taught.
- **Average Score** is the average of the students' observed scores.

Teacher growth measures are from SAS® EVAAS® multivariate, longitudinal analyses using multiple subjects, grades, and years of data for each student.
• **Average Percentile** shows where the Average Student Score falls in the distribution that includes the observed scores of all students in the state that tested in this grade and subject.

• **Average Predicted Score** is the average of the predicted scores for the teacher's students in this grade and subject.

• **Average Predicted Percentile** shows where the Average Predicted Score falls in the distribution that includes the predicted scores of all students in the state that tested in this grade and subject.