

Technical Description: Multi-Year and Composite Calculations

Introduction

This document captures how the policy decisions by the Ohio Department of Education (ODE) are implemented in the calculation of composites and multi-year trends for teachers in the tested subjects and/or grades.

While the following text provides a specific example of a teacher’s composite, the key policy decisions can be summarized as follows:

- A multi-year trend is calculated for an individual subject and grade for up to three years.
- A composite is calculated for multiple subjects and grades for up to three years.
- The composite for teachers includes only the subjects for which the teacher has a value-added measure in the current year.
- The composite for teachers weights each subject/grade/year equally.
- The composite for teachers uses the *most appropriate and robust* statistical approach possible in the calculation of the value-added estimate and associated standard error.

The composite for teachers will include OAA math and reading. The following examples will be used to show how the OAA multi-year trend in a single subject and the OAA composite across subjects would be calculated for a sample teacher.

Example 1: Available Data for OAA Multi-Year Trend for a Sample Teacher in a Single Subject

Year	Subject	Grade	Value-Added Gain	Standard Error
2010	Math	8	4.50	1.60
2011	Math	8	3.80	1.50

Example 2: Available Data for OAA Multi-Year Composite for a Sample Teacher across Subjects

Year	Subject	Grade	Value-Added Gain	Standard Error
2009	Science	8	4.20	2.00
2009	Math	7	3.50	1.50
2010	Reading	8	0.50	1.40
2010	Math	8	4.50	1.60
2011	Reading	8	-0.30	1.20
2011	Math	8	3.80	1.50

Calculating Gains for the OAA Multi-Year Trend and OAA Composite

For the teacher in Example 1, a multi-year trend can be calculated using the two years of data for this teacher in a specific subject and grade, 8th grade math. The multiple year trends in OAA math and reading will use re-estimated value-added gains and standard errors for years prior to the current year. This re-estimation will take into account current year student-level information to provide the most precise and reliable estimate of the prior year using all available information for that teacher in the year being analyzed. Each year used in the OAA multi-year trend is weighted equally, which ensures that teachers are neither advantaged nor disadvantaged due to one particularly different year. Because each group of students and each scenario are different every year, this approach will dampen any year to year variability. Because the value-added estimates are in the same scale (Normal Curve Equivalents), the composite gain across the years is a simple mean gain using all of the cells with equal weights from above.

The multi-year gain for Example 1 is calculated as follows:

$$Multi_{year}Gain = \frac{1}{2}Math_{8_{2010}} + \frac{1}{2}Math_{8_{2011}} = \frac{1}{2}4.50 + \frac{1}{2}3.80 = 4.15$$

For the teacher in Example 2, a composite gain that includes more than one subject would be calculated. A teacher's composite only includes the subjects for which there is a value-added report in the most recent year. As a result of this policy, the teacher is accountable only for the subject(s) that he or she currently teaches. There are a variety of reasons why a teacher may not teach a particular subject anymore, and this policy mitigates any concerns related to a deliberate decision by the teacher or his/her administrator to focus on other subject(s). As a consequence, this teacher's science report will be excluded from the composite since science had no value-added measure in 2011. Note that science would not be included in a math and reading composite regardless, but this illustrates the point of the subject being there in the current year. However, this teacher's 7th grade math report will be included, even though there was no value-added measure for 7th grade math in 2011, because there were value-added measures for the subject math in 2011. The last five rows of the chart above represent the five subject/grade/years that will be used in this sample teacher's composite.

Each subject/grade/year used in the OAA composite is weighted equally as was done with the years above. As in Example 1, because the value-added estimates are in the same scale (Normal Curve Equivalents), the composite gain across the five subject/grade/years is a simple mean gain using all of the cells with equal weights. The composite gain is calculated using the following formula:

$$\begin{aligned} Comp\ Gain &= \frac{1}{5}Math_{7_{2009}} + \frac{1}{5}Math_{8_{2010}} + \frac{1}{5}Math_{8_{2011}} + \frac{1}{5}Read_{8_{2010}} + \frac{1}{5}Read_{8_{2011}} \\ &= \frac{1}{5}3.50 + \frac{1}{5}4.50 + \frac{1}{5}3.80 + \frac{1}{5}0.50 - \frac{1}{5}0.30 = 2.40 \end{aligned}$$

Calculating Standard Errors for the OAA Multi-Year Trend and OAA Composite

The table for each of the two examples above reports a value-added gain as well as a standard error associated with that gain for each subject/grade/year. First of all, please note that the use of the word "error" does not indicate a mistake! Rather, value-added models produce *estimates*. That is, the value-added gains in the above tables are estimates, based on student test score data, of the teacher's true value-added effectiveness. In statistical terminology a "standard error" indicates the magnitude of the uncertainty in the estimate, providing a means to determine whether or not an estimate is decidedly above or below the growth expectation. Standard errors can, and should, also be provided for the multi-year and composite gains that have been calculated, as shown above, from a teacher's value-added gains.

A General Formula for the Standard Error of a Composite

First, a bit of terminology: the square of the standard error is called the variance, and it relates to estimation in the context of this document. Statistical formulas are often more conveniently expressed as variances. Standard errors of multi-year trends and composites can be calculated using variations of the general formula shown below. To maintain the generality of the formula, the individual estimates in the formula (think of them as value-added-gains) are simply called X , Y , and Z . If there were more or fewer than 3 estimates, the formula would change accordingly. Also to maintain generality, the composite is not limited to equally-weighted estimates, but remember that the OAA multi-year trends and composites *do* use equal weighting. Instead each estimate is multiplied by a different weight, a , b , or c .

$$\begin{aligned} \text{Var}(aX + bY + cZ) &= a^2\text{Var}(X) + b^2\text{Var}(Y) + c^2\text{Var}(Z) \\ &+ 2ab \text{Cov}(X, Y) + 2ac \text{Cov}(X, Z) + 2bc \text{Cov}(Y, Z). \end{aligned}$$

The “Cov” in the formula is the covariance. Covariance is a measure of the relationship between two variables. It is a function a more familiar measure of relationship, the correlation coefficient. Specifically, the term $\text{Cov}(X, Y)$ is calculated as follows:

$$\text{Cov}(X, Y) = \text{Correlation}(X, Y)\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}.$$

The value of the correlation ranges from -1 to +1, and these values have the following meanings.

- A value of zero indicates no relationship.
- A positive value indicates a positive relationship, or Y tends to be larger when X is larger.
- A negative value indicates a negative relationship, or Y tends to be *smaller* when X is larger.

Determining Statistical Independence

Two variables that are unrelated have a correlation, and covariance, of zero. Such variables are said to be statistically independent. This will be the case for multi-year trends as presented in Example 1. If the X and Y values have a positive relationship, then the covariance will also be positive. For the composite gains presented Example 2, the relationship will generally be positive, and this means that the OAA composite standard error is larger than it would be assuming independence.

Example 1: Standard Error and Index for OAA Multi-Year Value-Added Gain

As a general rule, two value-added gain estimates are statistically independent if they are based on completely different sets of students. This is almost always the case with multi-year trends in a single subject, which is the case in Example 1. It is unlikely, though not impossible, that any of this teacher’s 2010 8th grade math students were also in the teacher’s 2011 8th grade math class. With the assumption of independence, the formula for the standard error of the multi-year gain for Example 1 becomes fairly simple. Recall that the standard error is obtained by taking the square root of the variance.

$$\begin{aligned} \text{Multi}_{\text{year}}SE \text{ Gain} &= \sqrt{\left(\frac{1}{2}\right)^2 (SE \text{ Math}_{8_{2010}})^2 + \left(\frac{1}{2}\right)^2 (SE \text{ Math}_{8_{2011}})^2} \\ &= \frac{1}{2}\sqrt{(1.60)^2 + (1.50)^2} = 1.10 \end{aligned}$$

Using the multi-year value-added gain and multi-year standard error, it is possible to calculate an index. The index is simply the value-added gain divided by its standard error. In this example, the index is 4.15 divided by 1.10, which is 3.78 (using the unrounded multi-year standard error).

Example 2: Standard Error and Index for OAA Composite Value-Added Gain

Unlike the situation in Example 1, in this example the standard error of the OAA composite value-added gain cannot be calculated using the assumption that the gains making up the composite are independent. This is because it is much more likely that some of the same students are represented in different value-added gains, such as 8th grade math in 2011 and 8th grade reading in 2011. To demonstrate the impact of the covariance terms on the standard, it is useful to calculate the standard error using (inappropriately) the assumption of independence. The standard error would then be as follows:

$$\begin{aligned}
 SE \text{ Comp Gain} &= \\
 \frac{1}{5} \sqrt{(SE \text{ Math}_{7_{2009}})^2 + (SE \text{ Math}_{8_{2010}})^2 + (SE \text{ Math}_{8_{2011}})^2 + (SE \text{ Read}_{8_{2010}})^2 + (SE \text{ Read}_{8_{2010}})^2} \\
 &= \frac{1}{5} \sqrt{(1.50)^2 + (1.60)^2 + (1.50)^2 + (1.40)^2 + (1.20)^2} = 0.65
 \end{aligned}$$

At the other extreme, if the correlation between each pair of value-added gains had its maximum value of +1, the standard error would be as follows:

$$SE \text{ Comp Gain} = \frac{1}{5} \sqrt{\begin{aligned} &(1.50)^2 + (1.60)^2 + (1.50)^2 + (1.40)^2 + (1.20)^2 + \\ &2(1.5 * 1.6 + 1.5 * 1.5 + 1.5 * 1.4 + 1.5 * 1.2 + 1.6 * 1.5 + \\ &1.6 * 1.4 + 1.6 * 1.2 + 1.5 * 1.4 + 1.5 * 1.2 + 1.4 * 1.2) \end{aligned}} = 1.44$$

The actual standard error will fall somewhere between the two extreme values of 0.65 and 1.44 with the specific value depending on the values of the correlations between pairs of value-added gains. The magnitude of each correlation depends on the extent to which the same students are in both estimates for any two subject/grade/year estimates. For example, if the 2011 8th grade math and 2010 8th grade math classes had no students in common, then their correlation would be zero. On the other hand, if the 2011 8th grade math and 2011 8th grade reading classes contained many of the same students, there would be a positive correlation. However, even if those two classes had exactly the same students, the correlation would likely be considerably less than +1. The actual correlations and covariances themselves are obtained as part of the EVAAS modeling process. It would be impossible to obtain them outside of the modeling process. This process uses all of the information about which students are in which subject/grade/year for each teacher. While this approach uses a more sophisticated approach, it more accurately captures the potential relationships among teacher estimates and student scores. This will lead to the appropriate standard error that will typically be between these two extremes, which are 0.65 and 1.44 in this particular example.

As in Example 1, the final step is to express the composite value-added gain as an index, calculated by dividing the composite value-added gain by its standard error. In this example, the composite index for this teacher is 2.40 divided by a number between 0.65 and 1.44. If the actual standard error in this example were 0.75, then the index for this teacher would be 2.40 / 0.75 = 3.20.