Understanding the New Mexico A-F School Grading System

Module 2

Understanding the New Mexico Value Added Model

January, 2012

NMPED New Mexico Public Education Department

Hanna Skandera
Secretary-Designate, Public Education Department

Goals

To develop an accountability model that:

- Correctly holds schools accountable for student learning
- Captures important differences regarding achievement
- Avoids classifying schools based on characteristics outside their control
- Provides information for school improvement
- Creates the correct motivations for improvement

What Comprises a School Grade

A School's grade consists of three sets of factors:

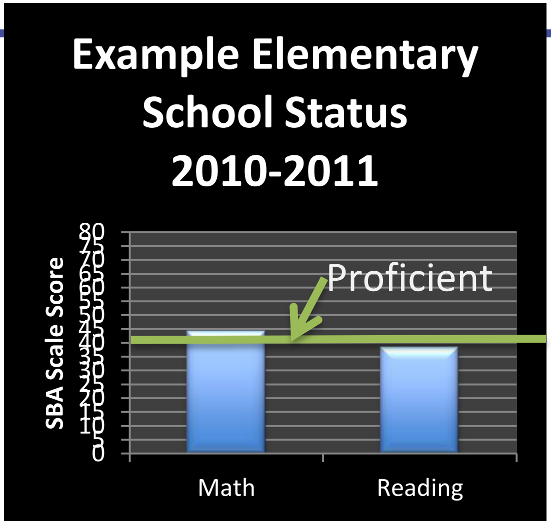
	Elementary and Middle Schools	High Schools
Current Standing	40%	30%
Growth	50%	30%
Opportunity to Learn	10%	8%
Graduation		17%
College and Career Readiness		15%

Current Standing

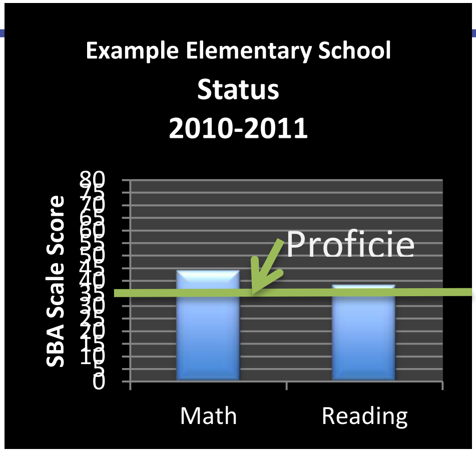
Current Standing consists of two components:

Percent Proficient

Conditional Status



Current Standing: Conditional Status



In thinking about holding schools accountable for student performance, we also need to acknowledge that schools serve different populations. We know that student performance is influenced by many factors and we want to isolate, as much as possible, what the school contributes to the students' scores.

This also "levels the playing field" by accounting for the different circumstances of students that schools face.

Current Standing: Conditional Status

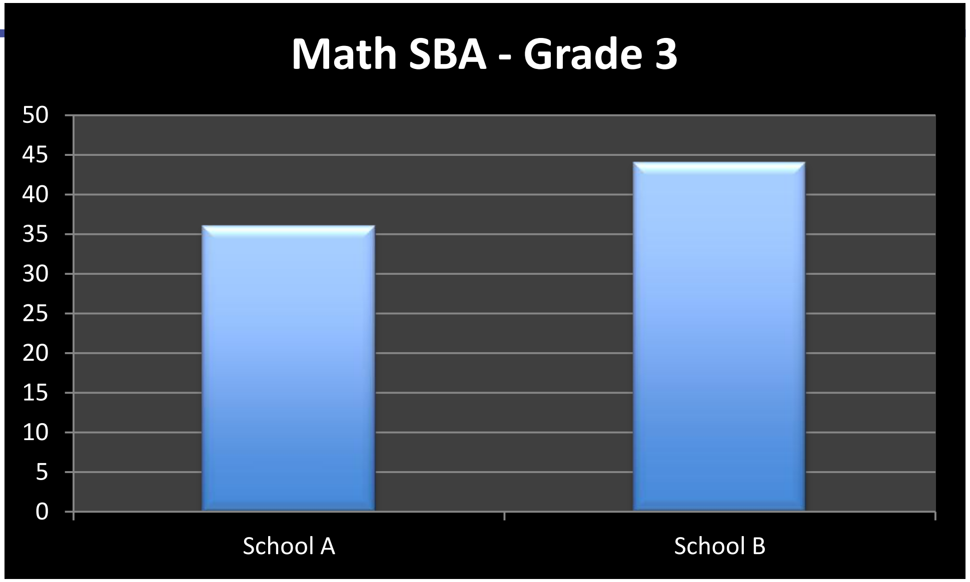
Including these student background variables to estimate the conditional status of a school in no way relates to different expectations for individual students. We expect every student to be college or career ready when they graduate high school.

To get a better understanding of what conditional information provides, we need to understand Value Added Models (VAM) New Mexico uses to estimate how schools are improving.

Value Added Models

We start with status and we can compare two schools, School A and School B (let's assume that School A and School B each have one student.)

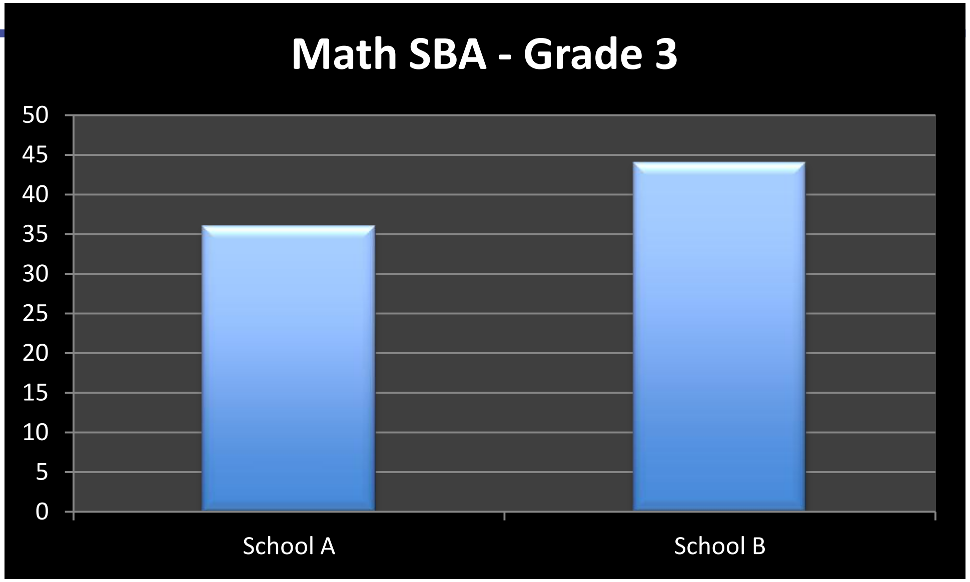
Based on the information presented in the chart, which school seems to be doing a better job teaching students?



Value Added Models

We start with status and we can compare two schools, School A and School B.

Based on the information presented in the chart, which school seems to be doing a better job teaching students?



School A has a score of 36, while School B has a score of 44. It would appear that School B is doing a better job.

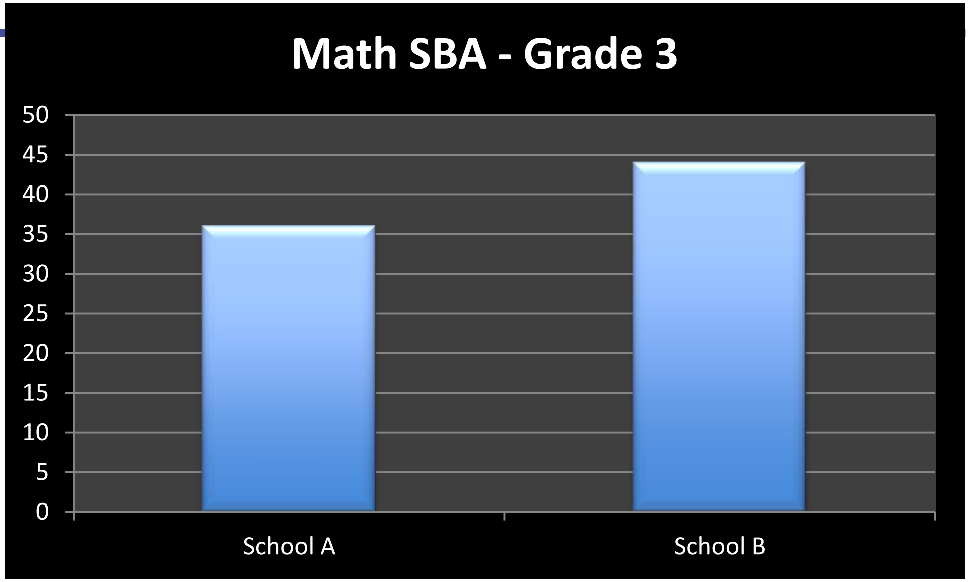
But, we want to be sure that we can attribute this success to School B.

Value Added Models

In order to unequivocally attribute student performance to a school, we would need to know that nothing is happening outside of the school that impacts a student's performance.

For example, if the student in School B received afterschool tutoring from *Bob's Afterschool Center*, then it may be that part of the success in School B is due to *Bob's Afterschool Center.*

If this is the case, should School B receive all the credit for that student's success?



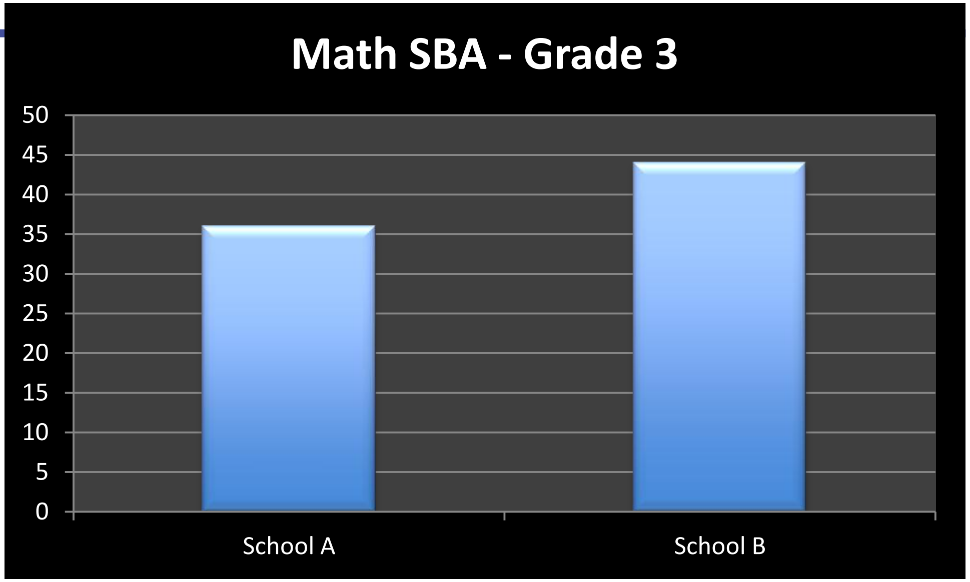
Value Added Models

If this is the case, should School B receive all the credit for that student's success?

Most would agree that School B should not get all the credit.

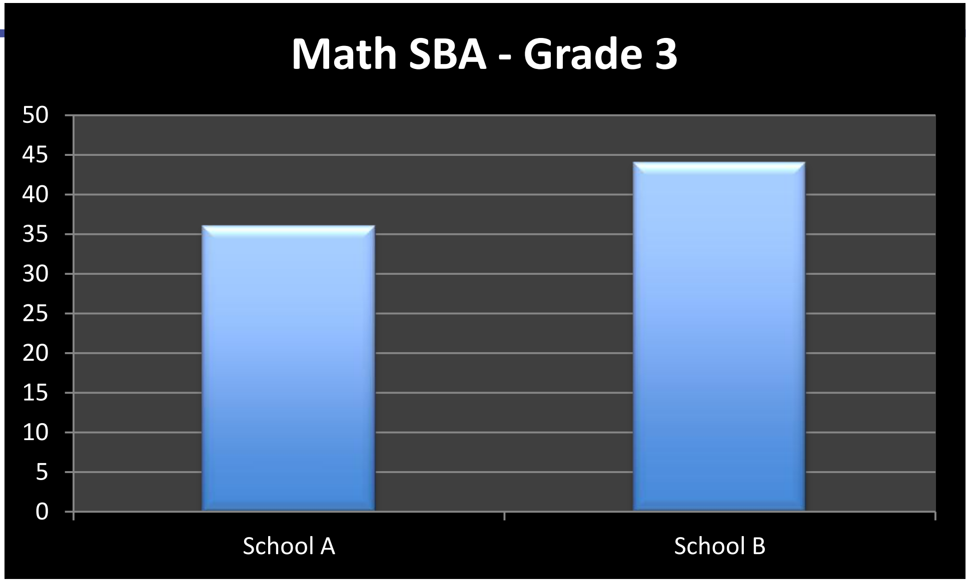
Bob's Afterschool Program is what we would call a factor that lies beyond a school's control.

We want to hold schools accountable only for factors that schools can control.



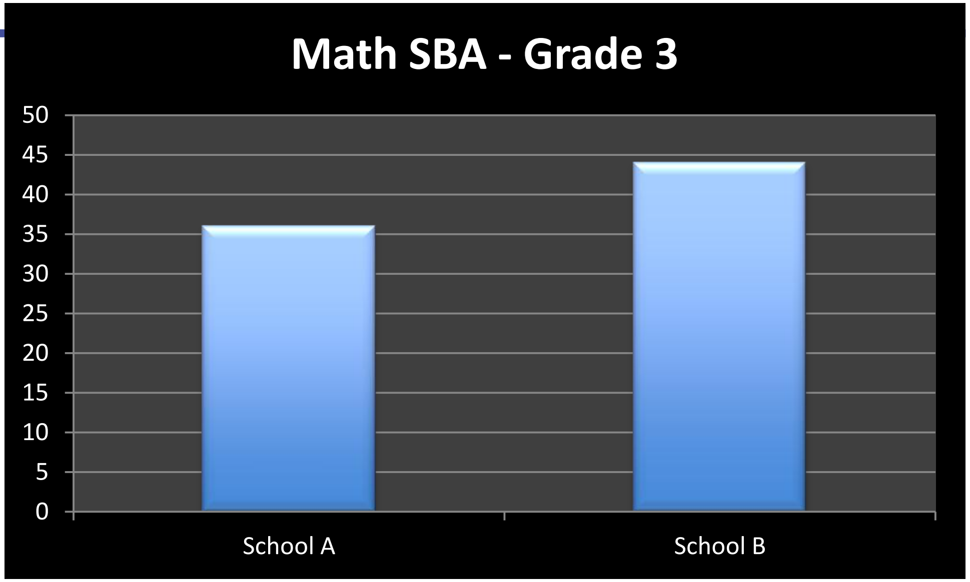
Value Added Models

The only way to assure that schools are fairly held accountable and don't receive extra credit for things they have no control over, or receive reduced credit for things they have no control over, is to assign students to schools randomly.



Value Added Models

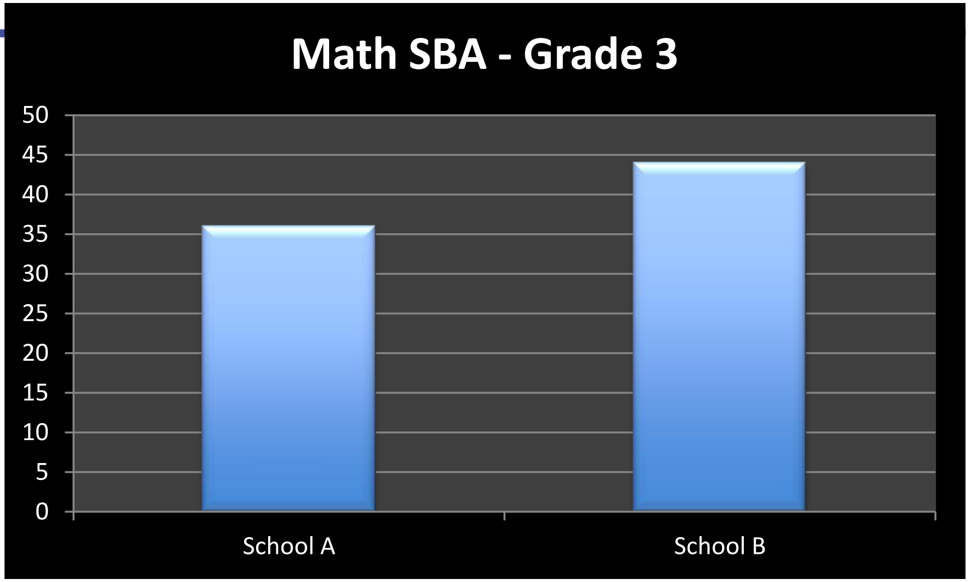
The only way to assure that schools are fairly held accountable and don't receive extra credit for things they have no control over, or receive reduced credit for things they have no control over, is to assign students to schools randomly.



Assigning students randomly to schools does not eliminate these uncontrollable factors from happening, it just assures us that they are equally likely to happen in all schools.

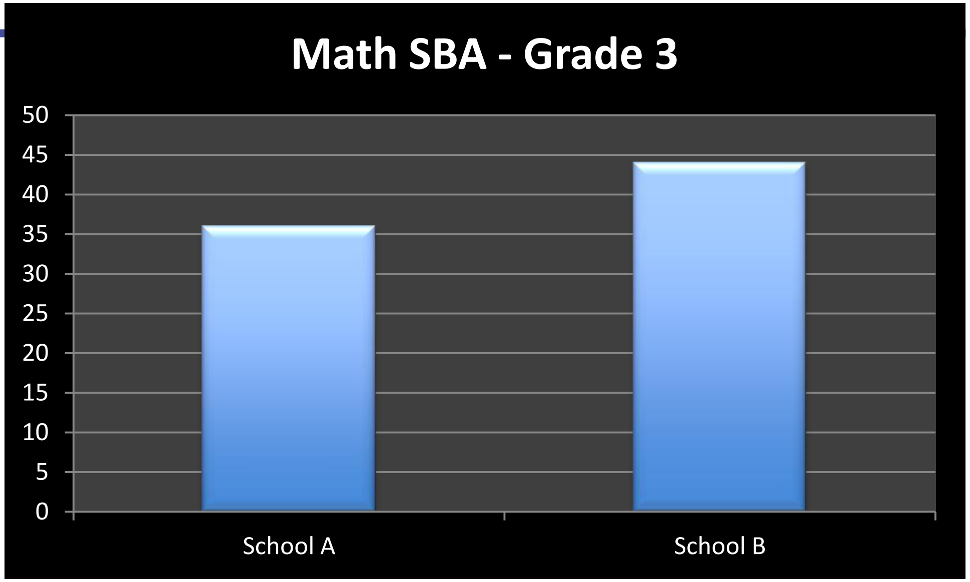
Value Added Models

It is not possible to assign students (and their teachers) to schools randomly, so we need to think of alternative mechanisms that help us isolate a school's contribution to student performance from what other non-controllable factors might contribute.



Value Added Models

It is not possible to assign students (and their teachers) to schools randomly, so we need to think of alternative mechanisms that help us isolate a school's contribution to student performance from what other non-controllable factors might contribute.



This can be done statistically. There are several approaches, but we focus on Value Added Models (VAM).

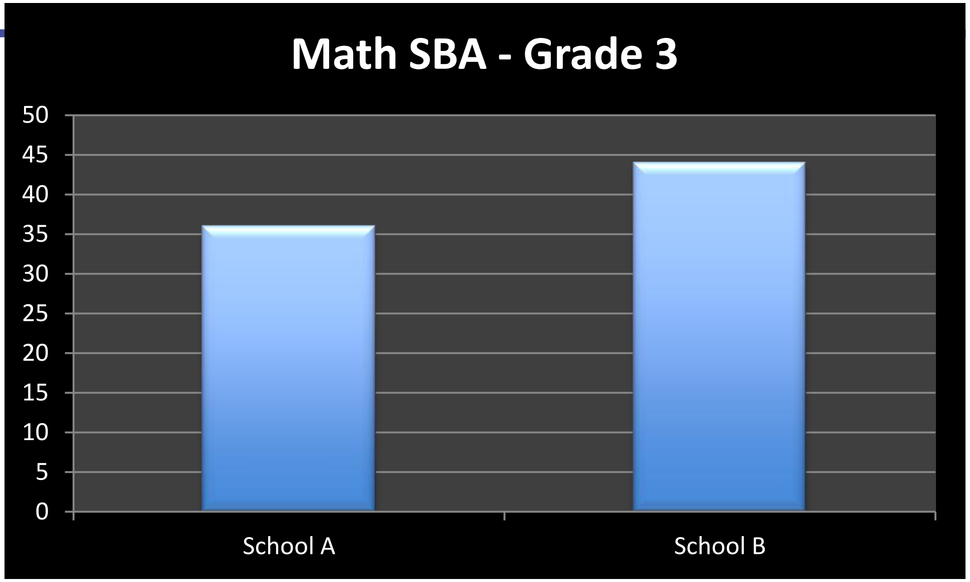
Even VAMs handle this differently, and what follows is the New Mexico model.

Value Added Models

One way of accounting for factors beyond school control is to predict how those factors might influence performance.

Unfortunately, we do not reliably and systematically collect data on every factor that might influence performance that is not within a school's control.

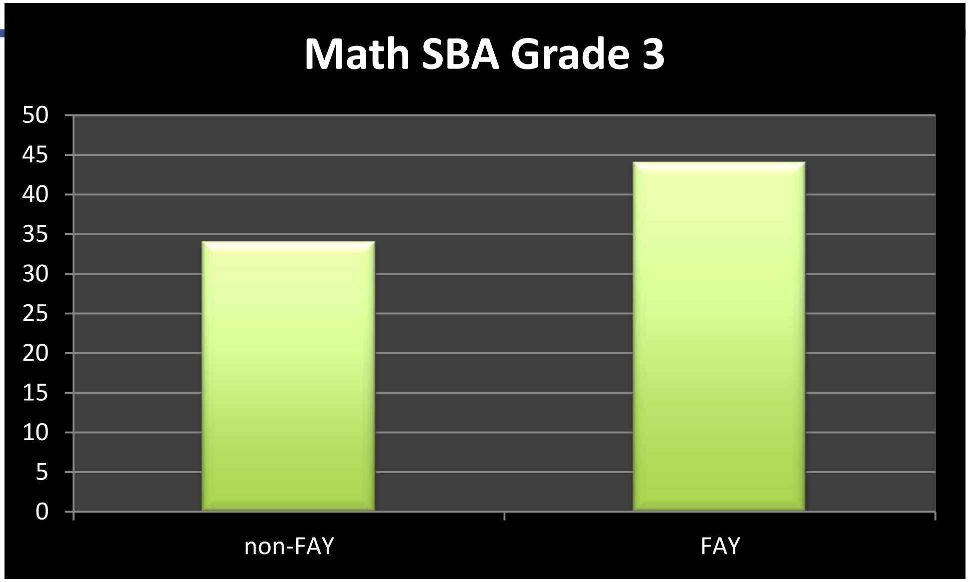
In fact, we collect very few – and often these are only proxies for what we might really want to know about a student.



Value Added Models

For example, we collect data on whether or not a student is Full Academic Year (FAY).

While we do not know exactly why not being FAY might matter – maybe the increased mobility disrupts students study habits, or the increased mobility does not give a student and teacher a chance to get to know each other and find the best ways to teach and learn.

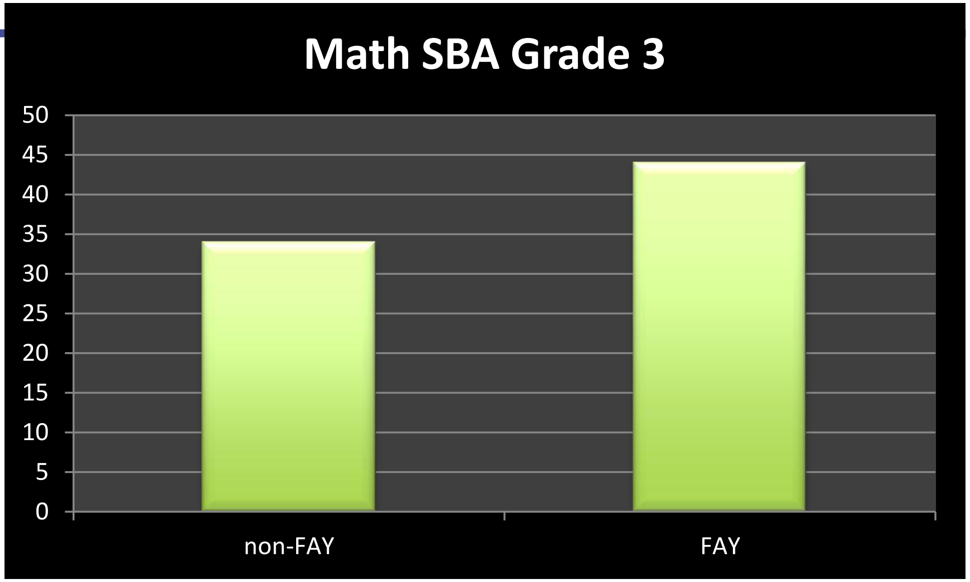


Value Added Models

For example, we collect data on whether on a student's Full Academic Year (FAY) status .

While we do not know exactly why not being FAY might matter – maybe the increased mobility disrupts students study habits, or the increased mobility does not give a student and teacher a chance to get to know each other and find the best ways to teach and learn.

We may not know why non-FAY students do not do as well, but we can observe that they do not do as well based on the data.

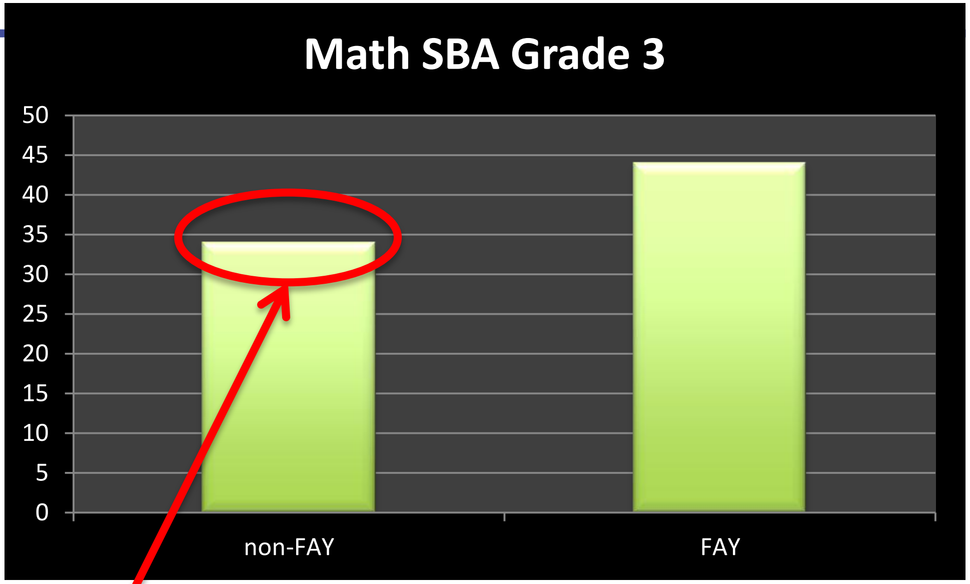


Value Added Models

For example, we collect data on whether on a student's Full Academic Year (FAY) status.

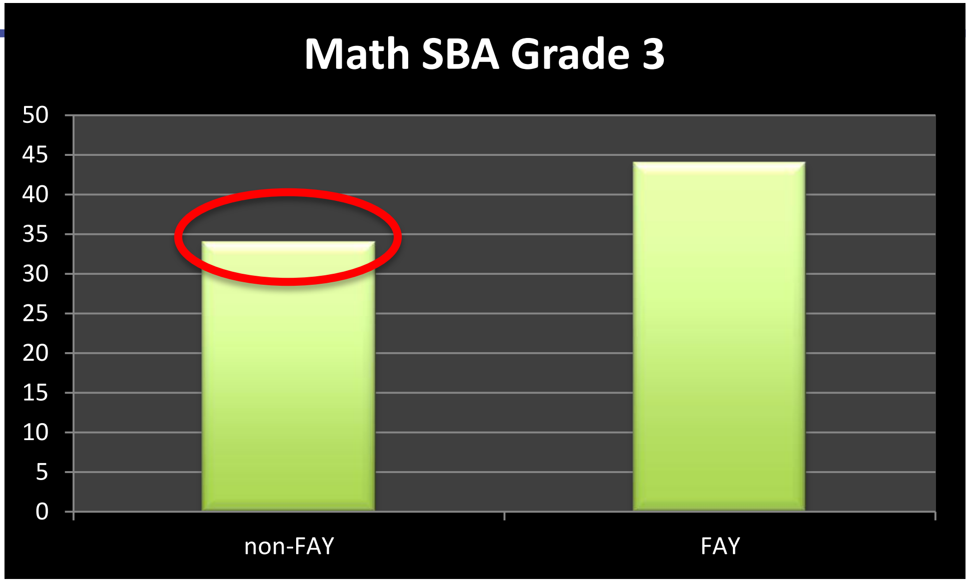
While we do not know exactly why not being FAY might matter – maybe the increased mobility disrupts students study habits, or the increased mobility does not give a student and teacher a chance to get to know each other and find the best ways to teach and learn….

We may not know why non-FAY students do not do as well, but we can observe that they do not do as well based on the data.



Value Added Models

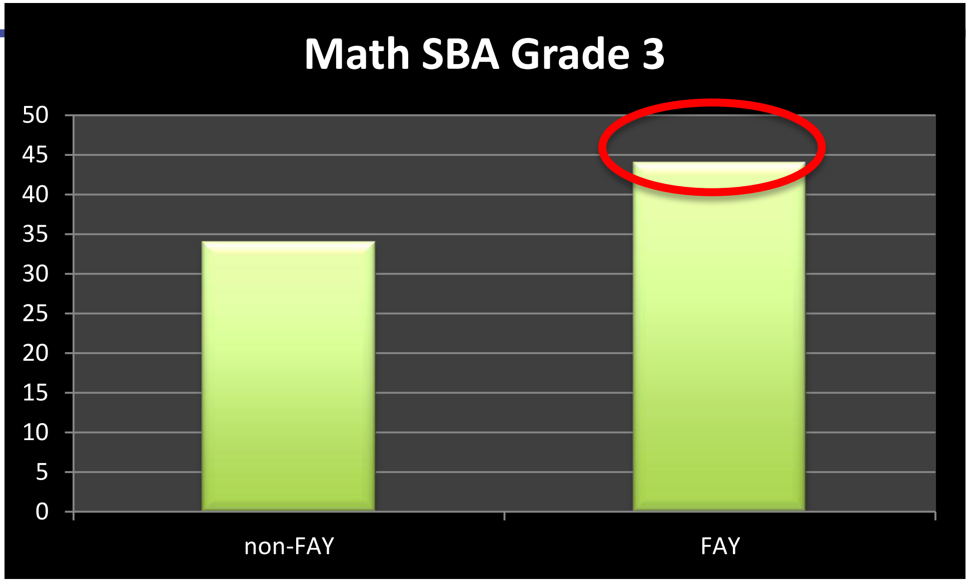
Given the data, we observe that non-FAY students score, on average, 34.



Value Added Models

Given the data we observe that non-FAY students score, on average, 34.

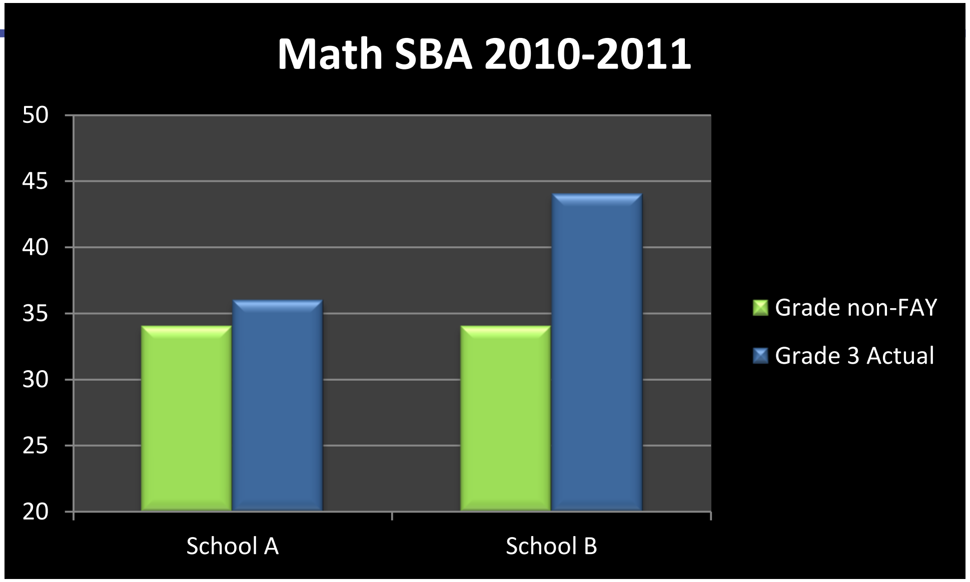
And that FAY students score, on average, 44.



Value Added Models

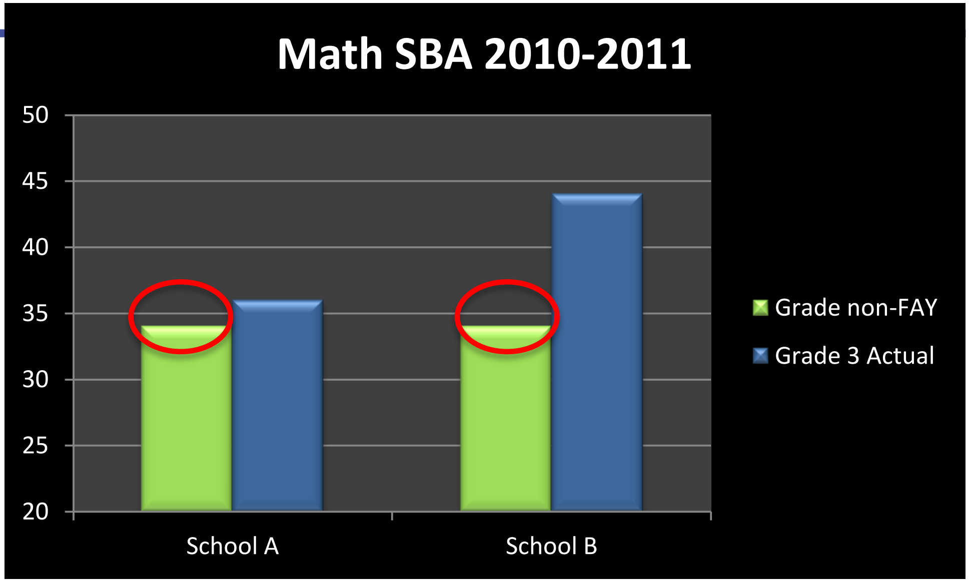
We can combine what we predict a student might score and compare this to what s/he actually scored.

Note: we have changed the vertical scale so it will be easier to make comparisons.



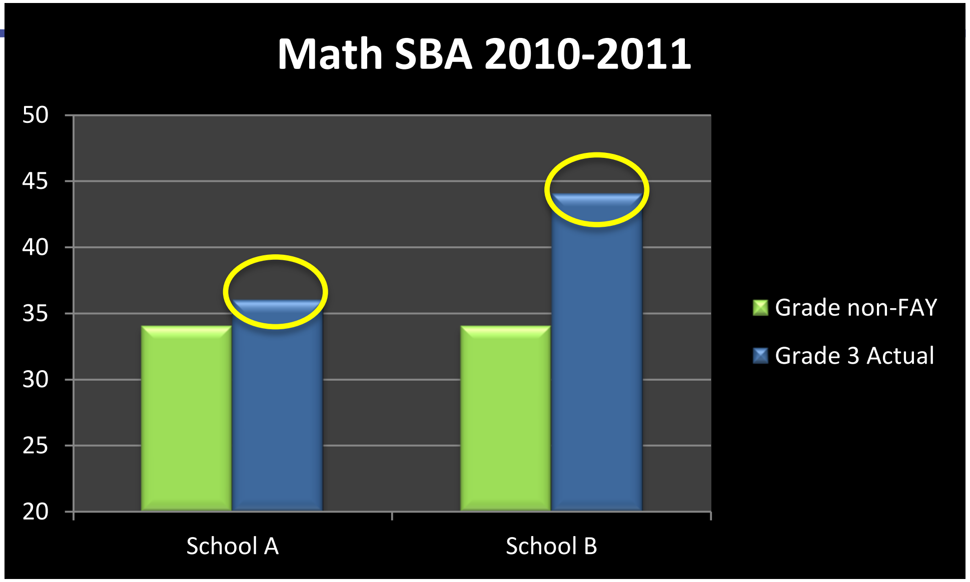
Value Added Models

Assuming both the student in School A and the student in School B are non-FAY, we can compare predicted performance.



Value Added Models

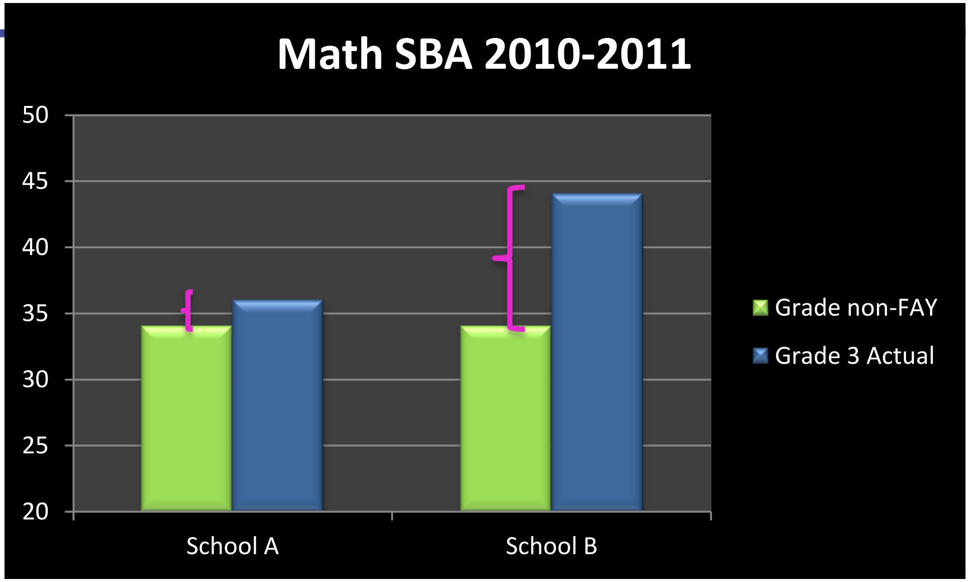
Assuming both the student in School A and the student in School B are non-FAY, we can compare predicted performance to the actual observed performance.



Value Added Models

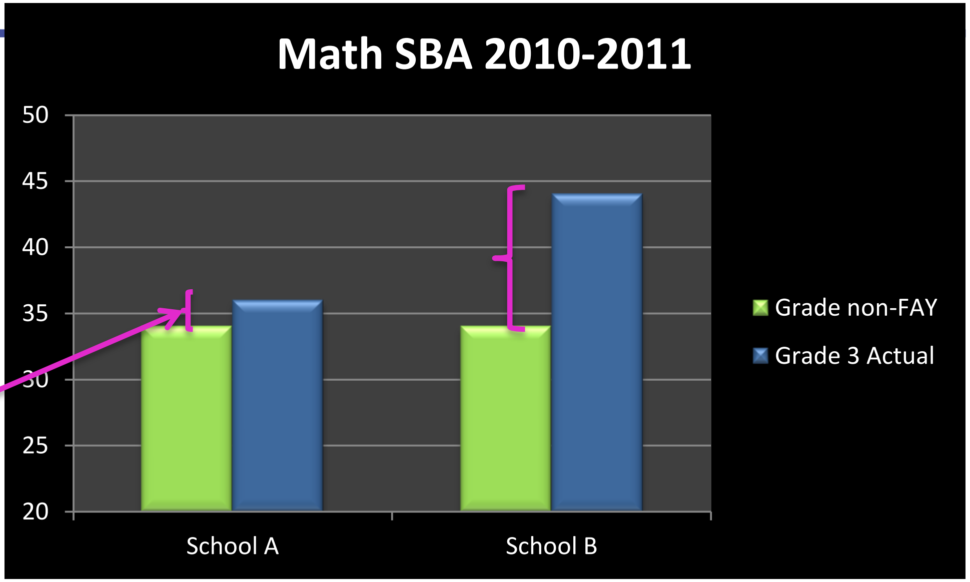
Assuming both the student in School A and the student in School B are non-FAY, we can compare predicted performance to the actual performance.

VAM score = Actual – Predicted



Value Added Models

Assuming both the student in School A and the student in School B are non-FAY, we can compare predicted performance to the actual observed performance.



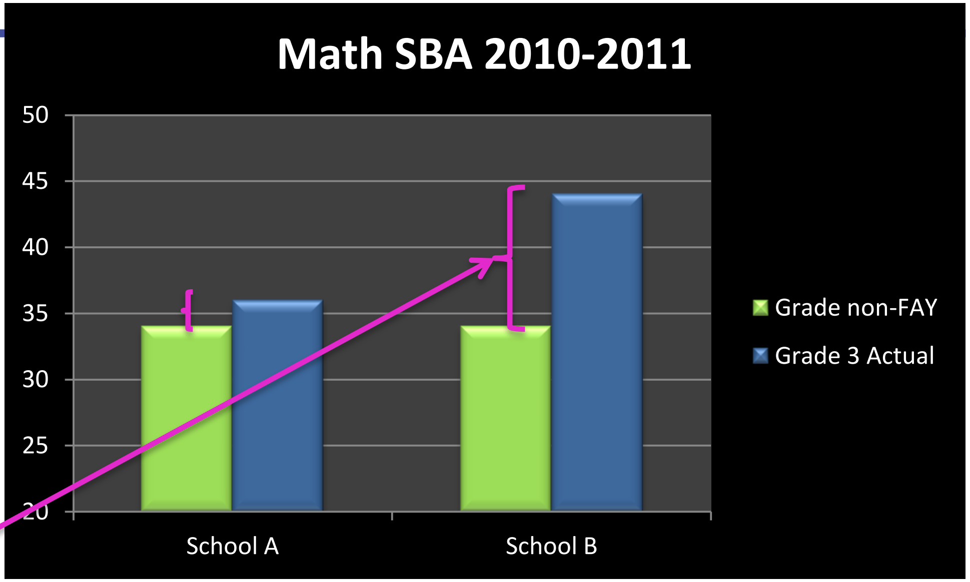
VAM score = Actual – Predicted

For School A

36-34 =2

Value Added Models

Assuming both the student in School A and the student in School B are non-FAY, we can compare predicted performance to the actual observed performance.



VAM score = Actual – Predicted

For School A

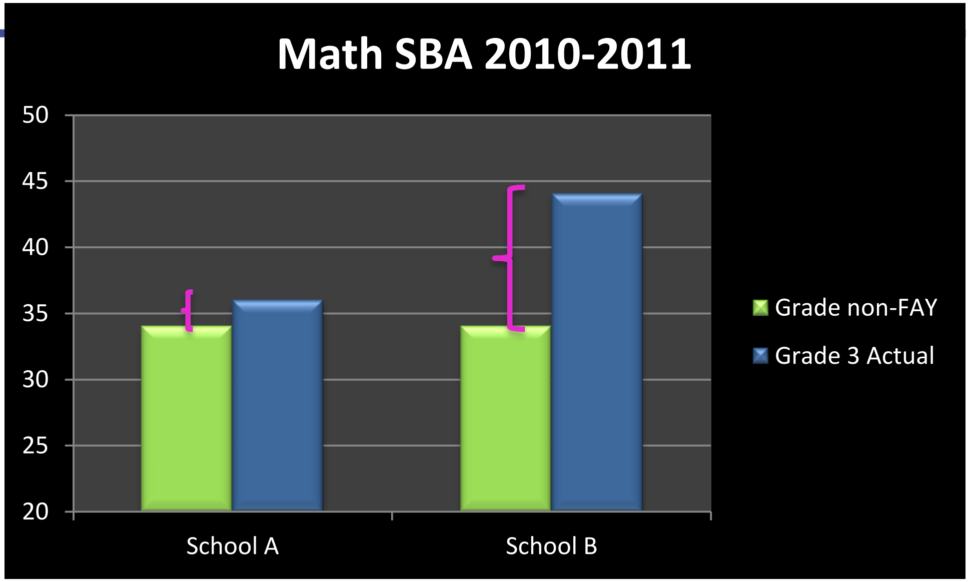
36-34 =2

For School B

44-34 =10

Value Added Models

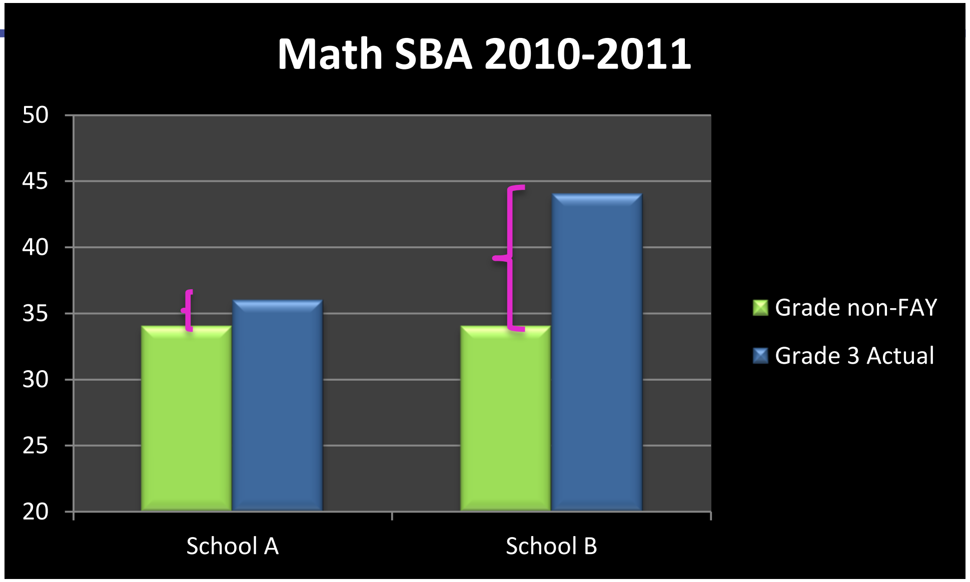
Using the VAM score does not change our initial interpretation of which school did a better job. School B has a VAM score of 10, while School A has a VAM score of 2.



Value Added Models

Using the VAM score does not change our initial interpretation of which school did a better job. School B has a VAM score of 10, while School A has a VAM score of 2.

If all schools had students with exactly the same background (e.g. non-FAY/FAY), then the VAM results would be exactly the same as simply looking at the actual scores.

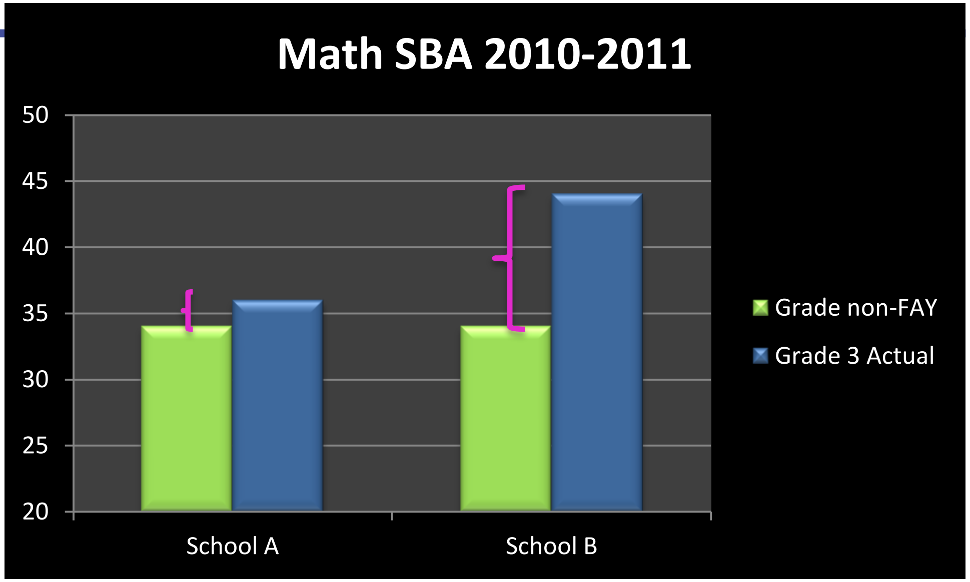


Value Added Models

Using the VAM score does not change our initial interpretation of which school did a better job. School B has a VAM score of 10, while School A has a VAM score of 2.

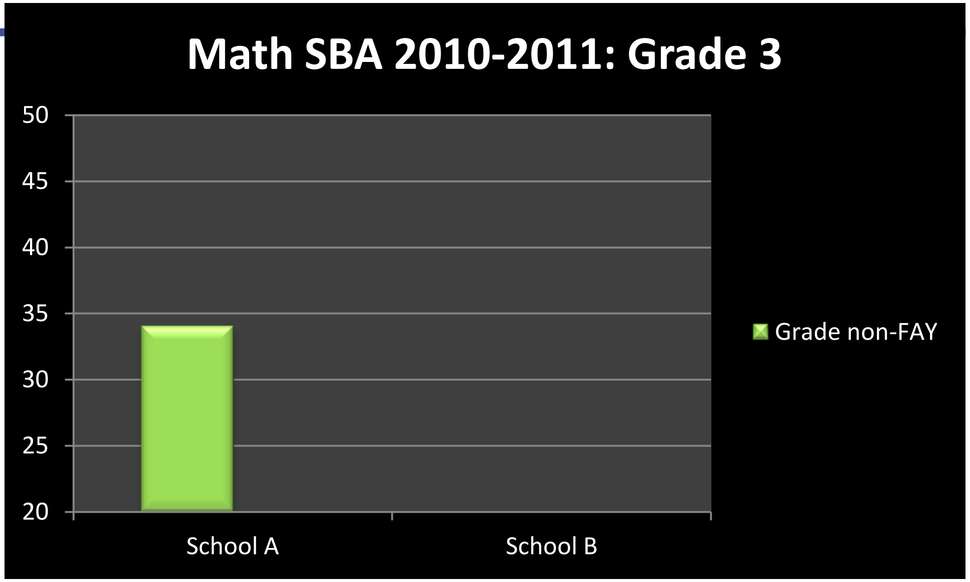
If all schools had students with exactly the same background (e.g. non-FAY/FAY), then the VAM results would be exactly the same as simply looking at the actual scores.

But, schools do not have students that are all the same, so VAM is still important to isolate effects for which schools ought not be held accountable.



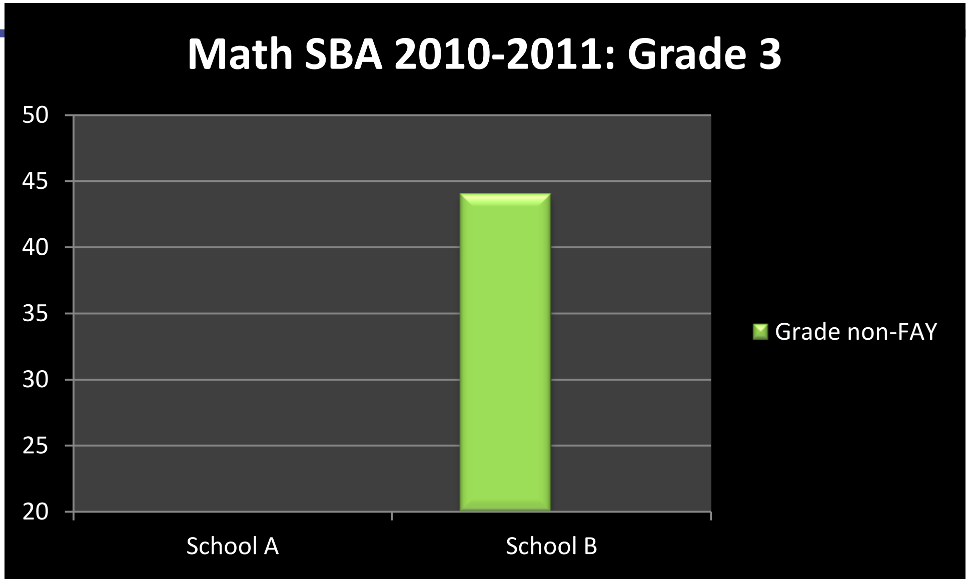
Value Added Models

Let us change the previous example just a little by assuming the student in School A is non-FAY and look at his predicted score.



Value Added Models

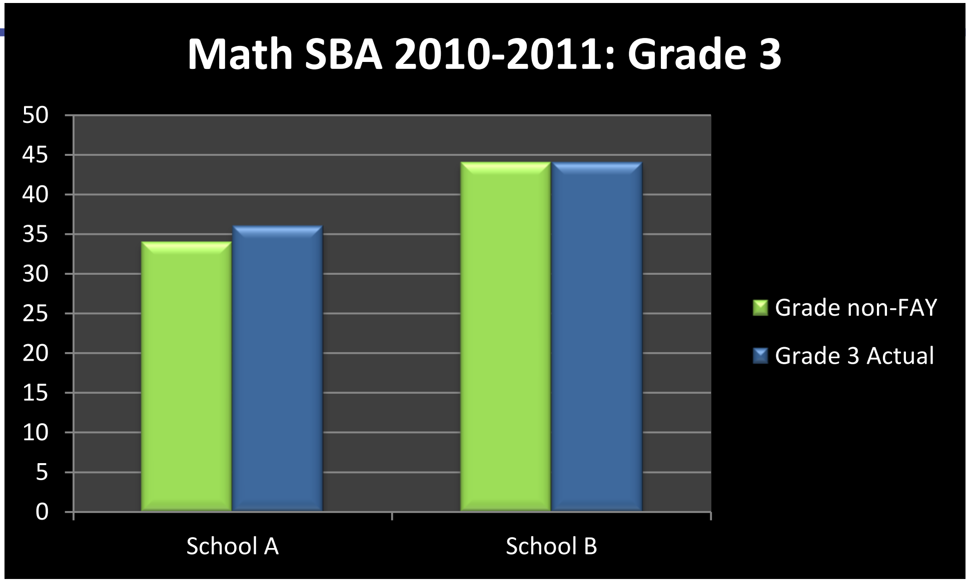
Let us change the previous example just a little by assuming the student in School A is non-FAY and the student in School B is FAY and look at her predicted score.



Value Added Models

Let us change the previous example just a little by assuming the student in School A is non-FAY and the student in School B is FAY.

We can then compare each student's predicted score to his or her actual observed score.



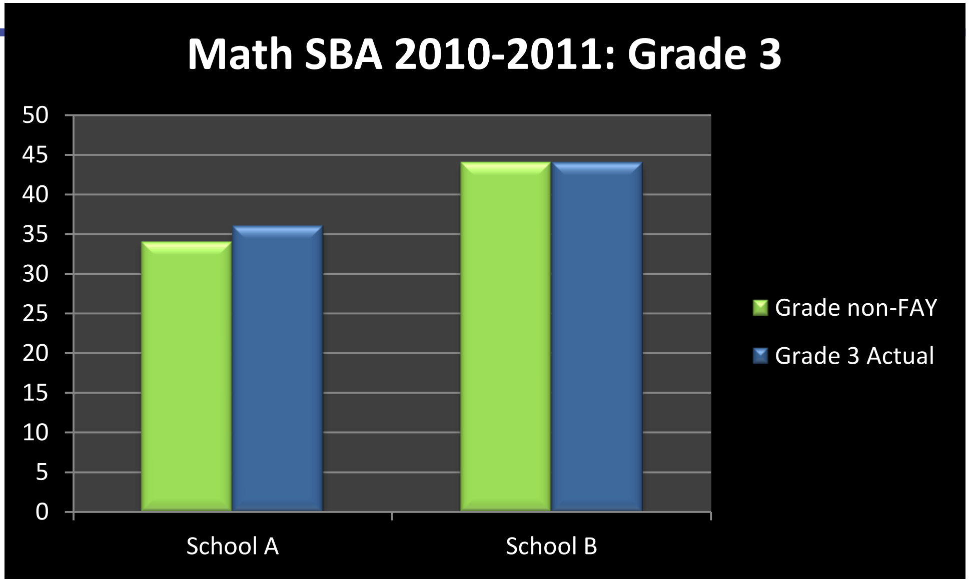
Value Added Models

Let us change the previous example just a little by assuming the student in School A is non-FAY and the student in School B is FAY.

We can then compare each student's predicted score to his or her actual observed score.

For School A
VAM score = 36-34 =2.

For School B
VAM Score = 44-44=0.



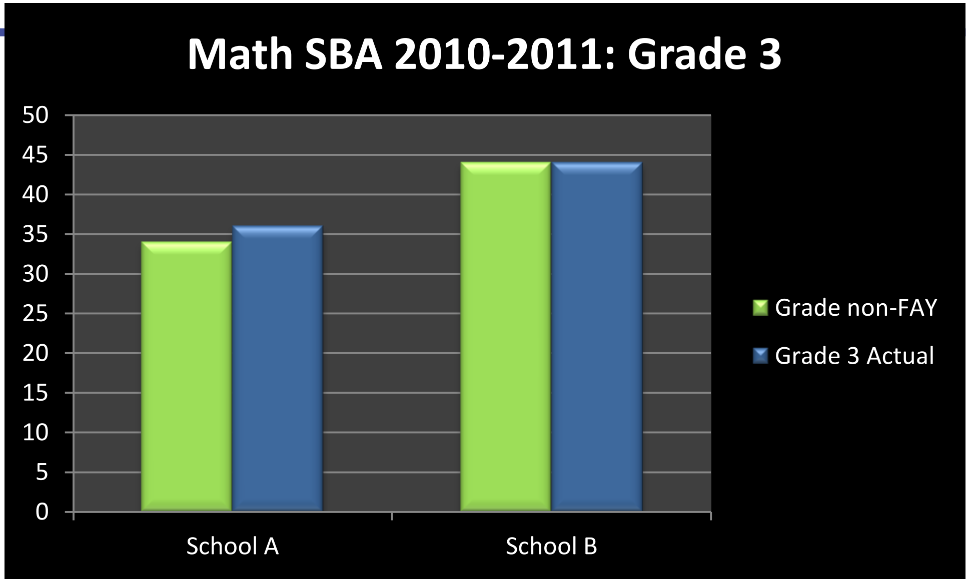
Value Added Models

For School A
VAM score = 36-34 =2.

For School B
VAM Score = 44-44=0.

We would now conclude, that accounting for the difference in how non-FAY students score compared to how FAY students score, we can attribute that portion of a student's success to the school.

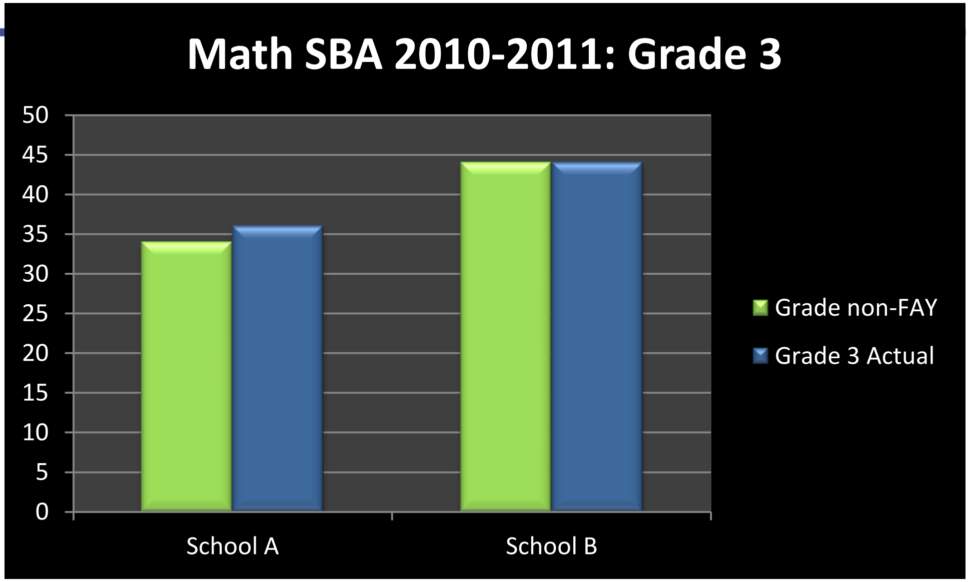
By dong this, we would conclude that School A is doing a better job than School B.



Value Added Models

Of course, FAY is not the only difference among students.

We include:
Gender
FAY
Language status
Economically disadvantaged (FRL)
Disability status (SWD)
Race/ethnicity
Grade
School size
Bottom quartile

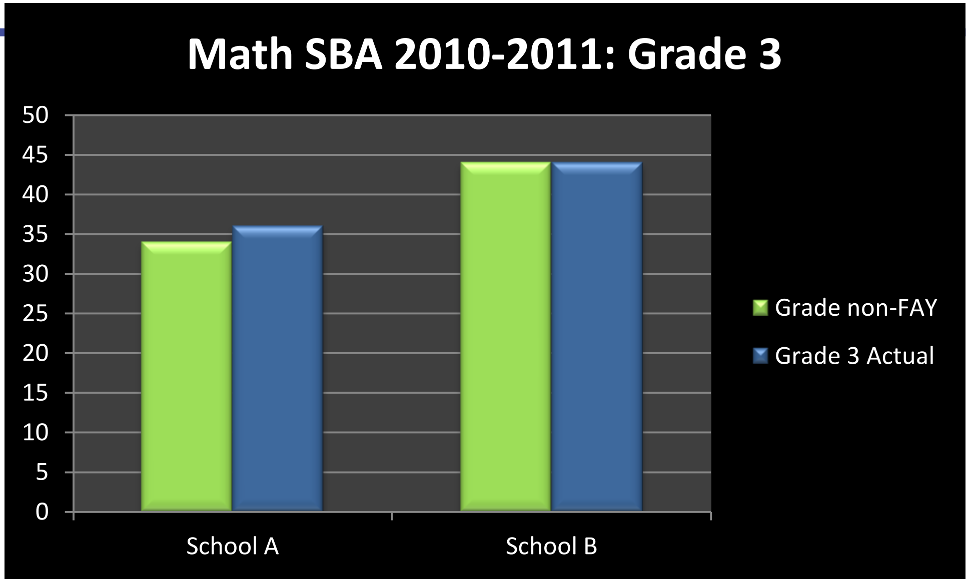


Value Added Models

Of course, FAY is not the only difference among students.

We include:
Gender
FAY
Language status
Economically Disadvantaged (FRL)
Disability status (SWD)
Race/ethnicity
Grade
School size
Bottom quartile

Let's consider on additional factor:
Economically Disadvantaged (FRL)



Value Added Models

Let's consider one additional factor:

Economically Disadvantaged (FRL)

We know that the predicted score for non-FAY is 34, and FAY is 44.

Value Added Models

Let's consider one additional factor:

Economically Disadvantaged (FRL)

We know that the predicted for non-FAY is 34, and FAY is 44.

It may be that FRL is 35 and non-FRL is 43.

Value Added Models

Let's consider one additional factor:

Economically Disadvantaged (FRL)

We know that the predicted for non-FAY is 34, and FAY is 44.

It may be that FRL is 35 and non-FRL is 43.

We only count each student once, so we also calculate the average for each of the combinations of FAY and FRL.

Value Added Models

Let’s consider on additional factor: Economically Disadvantaged (FRL)

We know that the predicted for non-FAY is 34, and FAY is 44.

It may be that FRL is 35 and Non-FRL is 43.

We only count each student once, so we also calculate the average for each of the combinations of FAY and FRL.

Predicted 3rd Grade Performance

	Non-Fay	FAY	Average
Not Econ Disadv.	38	48	43
Econ Disadv.	30	40	35
Average	34	44	

This example assumes equal cell sizes.

Value Added Models

A student who is both FRL and non-FAY is predicted to score 30.

<u>Predicted 3rd Grade Performance</u>

	Non-Fay	FAY	Average
Not Econ Disadv.	38	48	43
Econ Disadv.	30	40	35
Average	34	44	

This example assumes equal cell sizes.

Value Added Models

A student who is both FRL and non-FAY is predicted to score 30.

We would generate a prediction (again, based on actual average performance in 2010-2011) for every combination possible.

In this way, each student will have a prediction based on their individual characteristics.

Value Added Models

Of course, a school has more than one student and so the actual observed scores and predicted scores are calculated for each and averaged over all the students in the school.

Value Added Models

Of course, a school has more than one student and so the actual observed scores and predicted scores are calculated for each and averaged over all the students in the school.

In this example, the school has students that represent several different combinations of background characteristics.

Student	FAY	FRL	Predcited Score	Actual Score	Difference
1	N	Y	30	28	-2
2	N	Y	30	30	0
3	N	N	38	42	4
4	N	N	38	44	6
5	Y	Y	40	38	-2
6	Y	Y	40	40	0
7	Y	Y	40	44	4
8	Y	N	48	52	4
9	Y	N	48	50	2
10	Y	N	48	46	-2
Average					1.4

Value Added Models

Of course, a school has more than one student and so the actual observed scores and predicted scores are calculated for each and averaged over all the students in the school.

In this example, the school has the students and the students represent several different combinations of background characteristics.

The average of 1.4 can still be interpreted as the difference from predicted, but now on average for the students in that school.

Student	FAY	FRL	Predcited Score	Actual Score	Difference
1N	Y		30	28	-2
2N	Y		30	30	0
3N	N		38	42	4
4N	N		38	44	6
5Y	Y		40	38	-2
6Y	Y		40	40	0
7Y	Y		40	44	4
8Y	N		48	52	4
9Y	N		48	50	2
10Y	N		48	46	-2
Average					1.4

Value Added Models

We calculate the reliability of each school's estimate and make an adjustment based on the reliability.

(Note that school size is one factor that influences reliability).

In order to dampen the effect of unreliability (small school sizes), we create a "shrunken" VAM estimate that is shrunk towards the state average.

The amount we shrink each school's estimate depends on the reliability of the school's estimate.

Value Added Models

The average of 1.4 would be the school's VAM estimate for conditional status in math.

Because this 1.4 is an average, we can estimate the reliability of this score.

Reliability tells us how much of the variation in the observed scores is due to true variation and how much is due to error.

A reliability estimate of 1 indicates that all the variability in scores is true score variability

A reliability of 0 means the variation in scores is due to error.

Value Added Models

If the school's VAM estimate is perfectly reliable (=1.0), the "shrunken" estimate would equal the original VAM estimate.

As the reliability moves away from 1.0, we borrow information from the state average.

The formula works like this:

Value Added Models

If the school's VAM estimate is perfectly reliable (=1.0), the "shrunken" estimate would equal the original VAM estimate.

As the reliability moves away from 1.0, we borrow information from the state average.

The formula works like this:

Original VAM estimate X reliability + (1-reliability) X state average

Value Added Models

If the school's VAM estimate is perfectly reliable (=1.0), the "shrunken" estimate would equal the original VAM estimate.

As the reliability moves away from 1.0, we borrow information from the state average.

The formula works like this:

Original VAM estimate X reliability + (1-reliability) X state average

Our sample school has ten students and a VAM estimate of 1.4. The state average =0. Let's say the reliability of for this school is .80.

Value Added Models

If the school's VAM estimate is perfectly reliable (=1.0), the "shrunken" estimate would equal the original VAM estimate.

As the reliability moves away from 1.0, we borrow information from the state average.

The formula works like this:

Original VAM estimate X reliability + (1-reliability) X state average

Our sample school has ten students and a VAM estimate of 1.4. The state average =0. Let's say the reliability of for this school is .80, then:

1.4 X.80 + (1-.8) X 0=1.12

Value Added Models

We generate a score for each school in this way and we would place it on the distribution of scores to find the percentile rank as we described in Module 1.

Value Added Models

The previous example considered a single year of data, but the VAM model estimates both conditional status and school growth.

We address the school growth piece next.

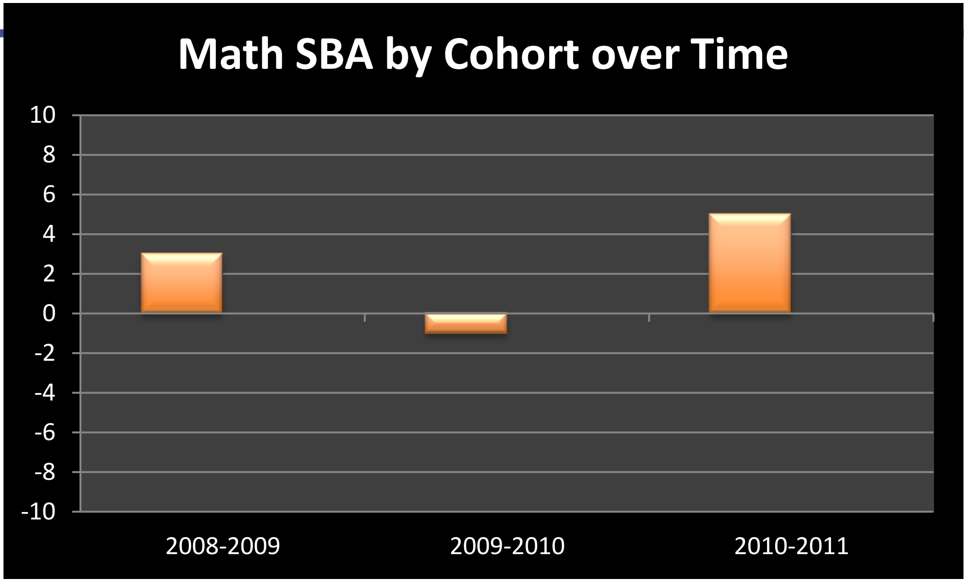
Value Added Models

Recall that the conditional status that we just estimated is based on the 2010-2011 characteristics of students and their scores.

We can go through this same exercise for any year.

In fact we do this for 2009 and 2010 and 2011 for the 2011 school grade.

Value Added Models



Recall that the conditional status that we just estimated is based on the 2010-2011 characteristics of students and their scores.

We can go through this same exercise for any year.

In fact we do this for 2009 and 2010 and 2011 for the 2011 school grade.

The chart demonstrates how actual performance is compared to predicted performance in each of the three years.

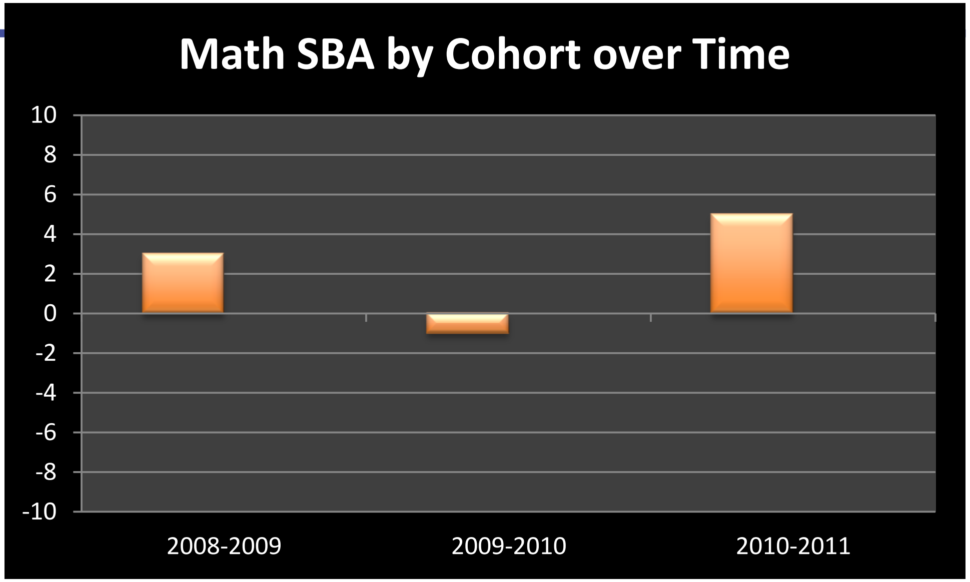
Value Added Models

Recall that the conditional status that we just estimated is based on the 2010-2011 characteristics of students and their scores.

We can go through this same exercise for any year. In fact we do this for 2009 and 2010 and 2011 for the 2011 school grade.

The chart demonstrates how actual performance is compared to predicted performance in each of the three years.

In 2008-2009 the school did better than predicted and in 2009-2010 it did a little worse.

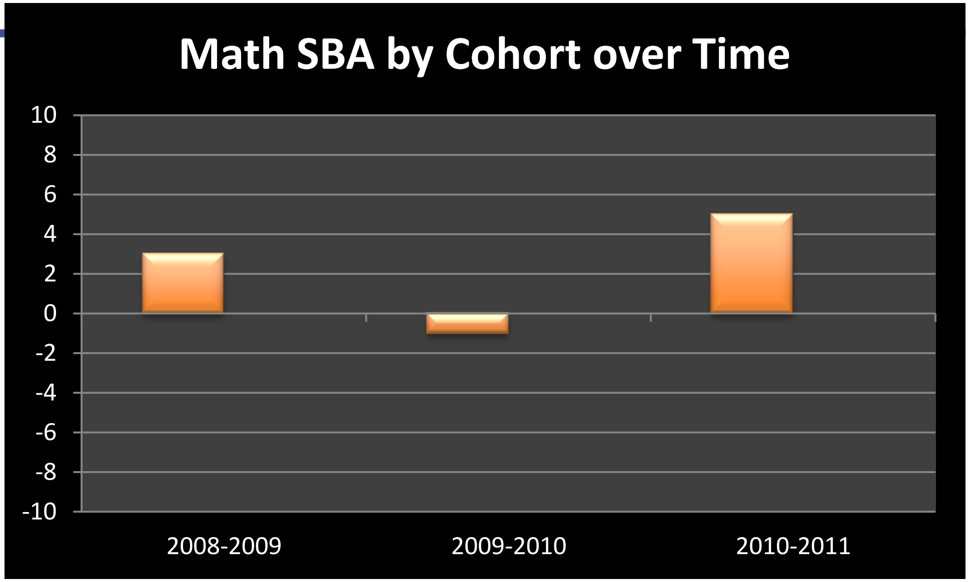


Value Added Models

We use the thee estimates and then see whether the variation over the years is related to time.

In other words, is there a trend?

On average is the school improving, or, as we term it, demonstrating school growth?

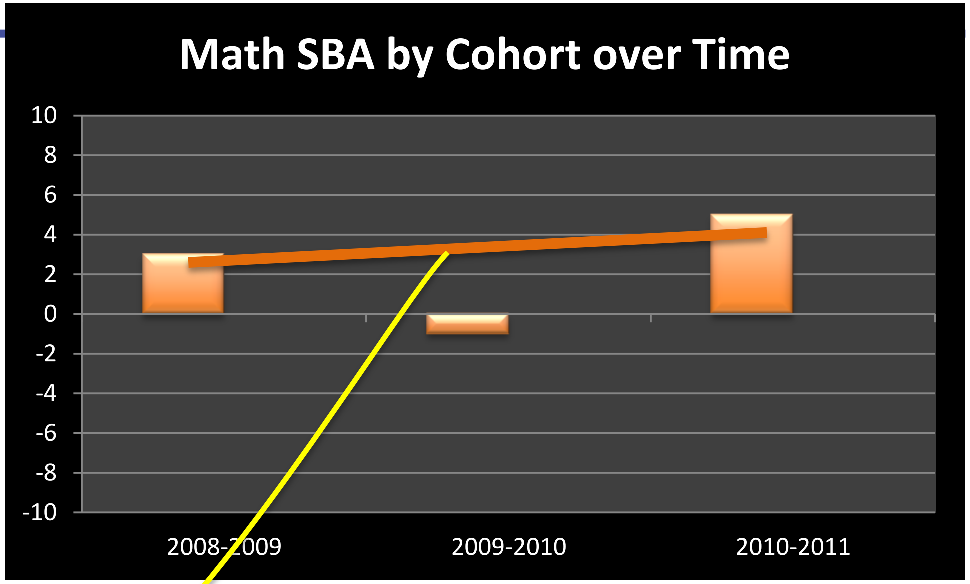


Value Added Models

We use the thee estimates and then see whether the variation over the years is related to time.

In other words, is there a trend?

On average is the school improving, or, as we term it, demonstrating school growth?



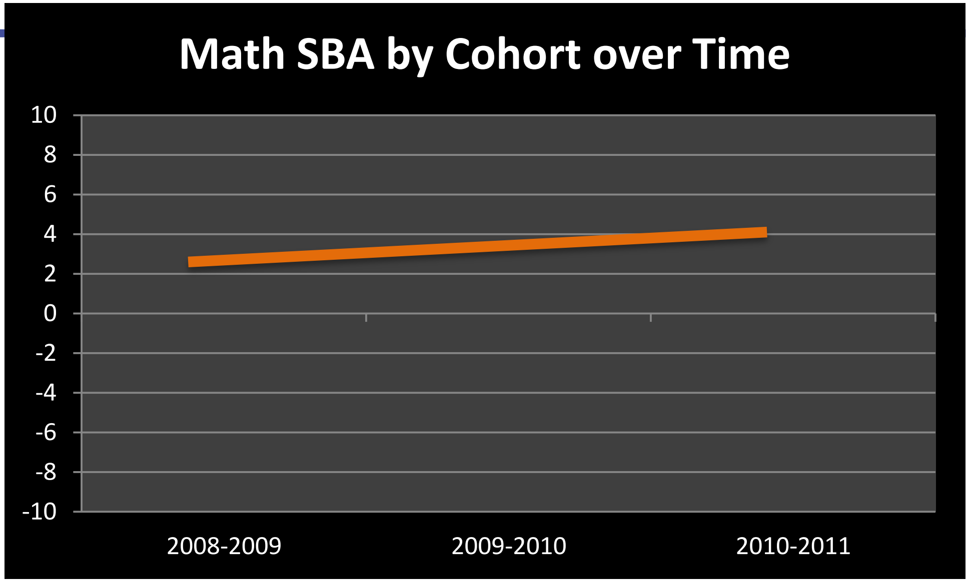
Value Added Models

We use the thee estimates and then see whether the variation over the years is related to time.

In other words, is there a trend? On average is the school improving, or, as we term it, demonstrating school growth?

We would take this estimate (about 1.5 points per year) and:

1) Calculate the "shrunken" score
2) Place this shrunken score on the distribution of all schools in the state
3) Find the percentile rank
4) Calculate points



Value Added Models

The New Mexico VAM estimates both a current conditional status and a trend.