Illinois Accountability Technical Advisory Committee (TAC)

Academic Growth for Accountability

Center for Assessment

December 18-19, 2017

Questions for the TAC

- What model features and characteristics are most important?
- What evidence should be collected to evaluate that the model is working as intended?
- What advice does the TAC have to ensure the ISBE's approach is feasible and minimizes operational risk and burden?
- What are promising approaches for identifying growth performance expectations?

Understanding Growth and Growth Models*

- Status vs. Improvement vs. Growth
  - **Status**: The academic performance of a student or collection of students at a single point in time (the single red dotted cell).
  - **Improvement**: Examining the change in performance over time within grades or across grades, without following the same student or collection of students (the gray vertical column and horizontal row).
  - **Growth**: The academic performance of *the same* student or *same collection* of students over two or more points in time (the diagonal green striped cells).

	Year					
Grade	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
3	**320**	**380**	**350**	**400**	**390**	**420**
4	**400**	**450**	**420**	**450**	**480**	**500**
5	**510**	**550**	**600**	**650**	**620**	**620**
6	**610**	**620**	**630**	**620**	**650**	**660**
7	**710**	**780**	**750**	**750**	**800**	**800**
8	**810**	**810**	**820**	**820**	**810**	**840**

*From a CCSSO presentation by D'Brot, J. & Goldschmidt,P., 2017

Common Growth Models*

*From Castellano and Ho, 2013

1. Gain Score (vertical scale score change)
2. Trajectory (average gains over time)
3. Categorical (transition tables)
4. Residual gain (difference between current and expected given past scores using regression)
5. Projection (past scores predicting future scores through regressions)
6. Student Growth Percentiles (percentile ranks, SGPs)
7. Multivariate (value-added models, e.g. VAAS)

Identifying Growth Models*

*From Castellano and Ho, 2013

*Adapted from Castellano and Ho (2013)

Generally Increasing Complexity

Primary Inference			
Statistical Foundation (classification)	**Growth Description**	**Growth Prediction**	**Value-Added**
Gain-Based Model: Based on score gains and trajectories on a vertical scale over time	1. Gain Score 3. Categorical Model	2. Trajectory 3. Categorical Model	1. Gain Score used as proxy for outcomes (i.e., links between aggregate/average gains and classroom/school membership)
Conditional Status Model: Scores in terms of expectations based on past scores	4. Residual Gain 6. Student Growth Percentiles	5. Projection Model 6. Student Growth Percentiles	4. Residual Gains using Covariate Adjustment (i.e., links between average conditional status and classroom/school membership)
Multivariate Model: Uses entire student score histories as an outcome to associate higher than expected scores with particular educators	Not typically applicable	Not typically applicable	7. Multivariate Model

Growth Model Selection: Key Considerations

- The theory of action driving the inclusion of growth in the accountability model.
- The type of inference you want (or need) to make about growth to support the TOA and intended use of reported growth measures.
- Required data and technical features
- Resource and logistical factors
- Stakeholder priorities

Growth Model Selection for School Accountability

What is the theory of action related to the inclusion of growth in the accountability model ?

- What policy goal is the inclusion of growth intended to support?
- What is the intended role of growth in the accountability model?
- What signal is the inclusion of growth in the model intended to send about the state's values and priorities?
- What factors should serve to differentiate schools with respect to growth?

Growth Model Selection for School Accountability

What type of interpretation does the growth measure need to support?

- Descriptive – interpretation of growth focused on observed achievement
- Predictive – interpretation of growth focused on future achievement
- Normative – interpretation of growth relative to the performance of others
- Criterion Referenced – interpretation of growth relative to a defined standard

Depends on:

- How you want stakeholders to use and interpret the growth measure.
- What is necessary to support the state's policy goals and theory of action.

Interpretations of Growth*

*Adapted from: Martineau, J., 2016: A Guide to Understanding and Selecting Measures of Growth

	Descriptive – Interpretations focused on observed achievement
Scale-referenced	The measure of growth is expressed as the number of points the student gained on the score scale.
Norm-referenced	Growth is expressed relative to the performance of a defined norm group.
Criterion-referenced	Growth is expressed relative to a defined expectation or standard
	Predictive– Interpretation focused on future achievement
Criterion-referenced: trajectory based	Continuation: Predict future performance based on the continuation of an observed trajectory of performance into the future. Target: Predict the trajectory of scores necessary to meet a particular target for performance in the future.
Criterion referenced: projection-based	Projected category: Predict whether a student will score in a desirable performance level in the future based on prior scores. Projected probability: predict the probability that the student will score in a desirable performance level in the future based on prior scores.

Growth Models*

* Adapted from Castellano and Ho (2013)

Characteristic	1. Gain Score (Raw, Simple, Average)	2. Trajectory (Growth to Standard)	3. Categorical (Transition Matrix, Value Table)	4. Residual Gain (Percentile Rank of Residuals, Resid Difference, Covariate Model)	5. Projection (Regression Model, Prediction Model)	6. Student Growth Percentile (Conditional Status %ile, %ile growth trajectory)	7. Multivariate (EVAAS, Layered Model, Cross-classified Model)
Description	Describes growth with simple differences or average gains over time	Extends gains or average gains in a predictable, usually linear fashion into the future	Defines growth by transitions among status categories (e.g., Basic, Proficient, Advanced) over time	Describes growth as the difference between current status and expected status give past scores	Uses past scores to predict future scores through regression equations	Percentile rank of current status in a reference group of students with similar scores	Uses entire student score histories, including other subjects and teachers, to detect higher than expected student scores associated with particular teachers
Question Addressed	How much has a student learned on an absolute scale?	If this student continues on this trajectory, where is she likely to be in the future?	How has this student grown in terms of transitions through categories over time? In which category will she likely be in the future?	How much higher or lower has this student scored than expected given her past scores?	Given this student's past scores, and based on patterns of scores in the past, what is her predicted score in the future?	What is the percentile rank of a student compared to students with similar score histories? What is the minimum SGP a student must maintain to reach a target standard?	Is this teacher associated with higher scores for his or her students than expected given all available scores and other teacher effects?
Primary Inference	Growth Description	Growth Prediction	Growth Description and Growth Prediction	Growth Description	Growth Prediction	Growth Description and Growth Prediction	Value-Added
Statistical Basis	Gain-Based	Gain-Based	Gain-Based	Conditional Status	Conditional Status	Conditional Status	Multivariate
Required Data Features	Vertical scale	Vertical scale	Articulated cut scores across years and grades. Values for value tables. Implicit vertical scale.	An interpretable scale. Assumptions of linear regressions.	Interpretable future scale or future standard.	Large sample sizes for reliable estimation.	For high-stakes value-added, many years of student data for stable teacher effects.
Group-level Interpretation	Average gain	Average trajectory or percentage of on-track students	Average across value tables or percentages of on-track students	Average residual gain	Average future prediction or percentage of on-track students	Median or average SGP, percentage of on-track students	Only group-level interpretations: Teacher- and school-level "effects"

Additional Considerations *

* Adapted from: D'Brot, J., 2017 Considerations for including growth in ESSA State Accountability Systems

- Resources available to conduct the required analysis (monetary and human) and QC the results
  - Computational complexity
  - Cost
- Logistical constraints
  - Data collection/storage requirements
  - Time
- Stakeholder priorities
  - Ease of communication
  - Transparency
- Technical characteristics
  - Will the approach provide for meaningful differentiation among schools

State Overview: ASR Survey of 32 States

Growth Model	Count	Examples
Student Growth Percentiles	15	AL, GA, HI, IN, KY, MA, ME,MI, NV, OR, SD, UT, WA, WV, WY
Value Added	8	AR, LA, NC, NM, OH, PA, TN, WI
Predictive/ Trajectory	1	CT
State Developed Model	3	IA, MD, NE
To be determined	5	AK, AZ, DE, KS, ID

Note: 3 states indicated that they were using a "growth to standard" approach. (CT, IN, LA)

From IL State Plan

- "Based on comments from the Governor and Stakeholders growth received over two times as much weight as proficiency in the accountability system."
- "Illinois recognizes an emphasis on student growth as a primary driver to close equity gaps. As a result, student growth will represent 50% of the accountability framework for Illinois."
- "ISBE proposes to utilize linear regression (i.e., current test scores are regressed on last year's test scores) to compute student academic growth in grades 3-8."

Questions for the TAC

- What model features and characteristics are most important?
- What evidence should be collected to evaluate that the model is working as intended?
- What advice does the TAC have to ensure the ISBE's approach is feasible and minimizes operational risk and burden?
- What are promising approaches for identifying growth performance expectations?

Model Evaluation Framework

*Slide taken from a CCSSO presentation by D'Brot, J. & Goldschmidt, P., 2017

- Policy-based Criteria
  - **Policy Goals/ Purpose**: *Alignment with Theory of Action. Why include growth?*
  - **Interpretation/ Inference Supported**: *What does the model tell us? Does it extrapolate beyond observed data?*
  - **Equity**: *Does the model expand our view of school quality fairly?*
  - **Resistance to Corruptibility**: *How "gameable" are prior scores?*
  - **Cost:** *How much does this thing cost to run?*
  - **Utility**: *Is if useful? Does it support changes in behavior?*

- Technically-based Criteria
  - **Technical Goals/Purpose**: *How well does it differentiate individually and as part of the system?*
  - **Analytical Approach**: *What is the underlying computation for the model?*
  - **Data Availability**: *What are the data requirements of the model?*
  - **Consistency**: *What kind of stability does the model support in its classifications (e.g., are we factoring in scale drift)?*
  - **Equity**: *How unrelated are the model outputs to supposedly unrelated data (e.g., school and LEA demographics)*
  - **Complexity**: *How complex are the results to explain to the public? How easily can results be integrated into the accountability system?*