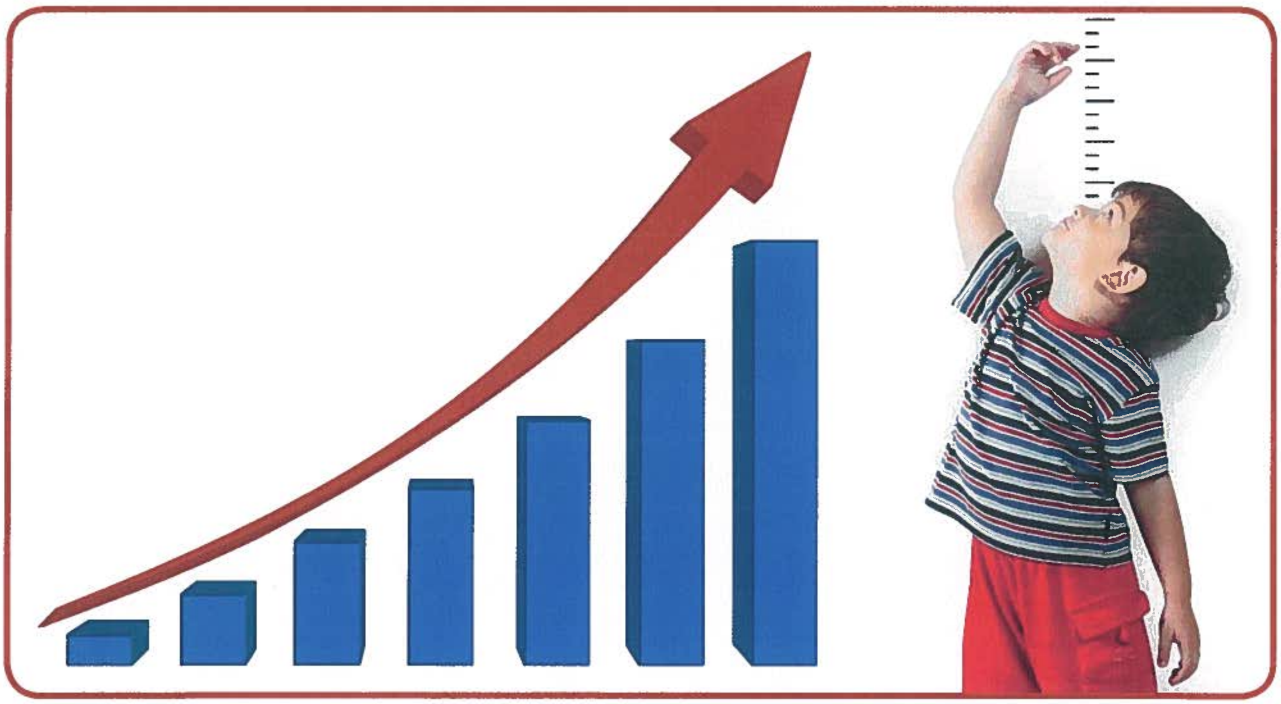
A Practitioner's Guide to
Growth Models



Katherine E. Castellano
University of California, Berkeley

Andrew D. Ho
Harvard Graduate School of Education

February 2013

6.2.1 Gain-based models

The first type of statistical foundation underlies models that are based on gains, average gains, or score trajectories over time. We call these *gain-based models*. A gain or gain score is the simple difference between two scores at different points in time. The gain score can be extrapolated over future time points to support predictions. When there are more than two data points for an individual, the gain can be generalized over multiple time points by averaging and expressing progress as an average change per unit of time.

A common feature to all gain-based models is an implicit or explicit recognition of a *vertical scale*, a common scale that allows scores to be compared across different grade-level tests. Vertical scales support interpretable score differences over the time and grade range of interest. A gain-based statistical foundation is consistent with an intuitive definition of growth: the difference between where one was and where one is. However, vertical scales are difficult to design and maintain, and many useful questions about performance over time do not require vertical scales. This motivates a contrasting statistical foundation underlying a second class of growth models.

6.2.2 Conditional status models

The second type of statistical foundation supports interpretations about *conditional status*. The word "conditional" implies an "if" statement, a kind of dependence, and, indeed, conditional status recasts or reframes status with respect to additional information. Models that use this statistical foundation address the question: How well does a student perform with respect to *expectations*? These expectations are set empirically using the past scores of the student of interest and other students.

Using this past information, conditional status models use a two-step process. First, given a student's past scores, they establish expectations about his or her current score. Second, the student's actual status is compared to these "conditional" expectations given past scores. The use and differentiation of past and current scores allows this method to meet our definition of a growth model. The phrase, "conditional status," is a technical term arising from the models' referencing of student status in terms that are conditional upon past scores or, more simply, in terms that consider past scores or take past scores into account. This foundation is fundamentally distinct from models that have a gain-based foundation, where status is evaluated over time instead of compared to expectations based on past scores.

Notably, conditional status models can reference current status to other variables in addition to or in place of past scores, such as economic status, race and ethnicity, or participation in specific educational programs. It is entirely possible to use a conditional status model to describe status in terms of expectations set by less relevant variables like a student's height or shoe size. This observation does not invalidate conditional status models as growth models but serves to emphasize how this statistical foundation supports a fundamentally different conception of growth: status with respect to expectations based on past scores and, potentially, other information.

A natural corollary of this definition of growth is that conditional status will change as expectations change. Setting expectations based upon two past scores will result in a different conditional status than setting expectations based on three past scores, and setting expectations based upon student demographic variables will change a student's conditional status score even further. In comparison, gain-based scores will also change under inclusion of additional time points. However, increasing previous time-points for *gain-based models* allows for better estimation of average gains, whereas using more past scores in *conditional status models* changes the substantive interpretation of the conditional status score. In sum, the output of conditional status models is interpreted most accurately with full appreciation of the variables that have been used to set expectations.

Conditional status scores can be reported on many metrics, from the test score scale to percentile ranks. As an example, consider a student whose high current status places her at the 80th percentile (among all students). In spite of this relatively high score, this student's past scores have been at even higher percentiles. Thus, her current percentile rank of 80 is somewhat below the empirically derived expectations given these past scores. One expression of conditional status is the simple difference between her actual current score and the score that is expected given her past scores. This describes the residual gain model in Chapter 4. Another approach expresses this low expectation in terms of a percentile rank. This latter approach is known as a Student Growth Percentile and is described in detail in Chapter 6. Table 1.5 displays conditional status models in its second row, cross-classified by the primary interpretations that these models support.

Chapters 4-6 review conditional status models and delve more deeply into the contrasts between gain-based and conditional status models. Understanding these contrasts is essential for accurate selection and use of growth models.

6.2.3 Multivariate models

The third type of statistical foundation is used primarily to estimate the "value-added" associated with classrooms and schools. Table 1.5 displays multivariate models in its third row and includes no models in the first two columns, as this statistical foundation is not well suited for growth description or growth prediction.

Multivariate models are distinguished by their complexity and their ability to use a large amount of data and variables in a unified approach. They require specialized and sometimes proprietary software and training in the interpretation of model output. The models are designed to

Summary Table

Model / Characteristics	Gain Score	Trajectory	Categorical	Residual Gain	Projection	Student Growth Percentile	Multivariate
Brief 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 given 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 past scores	Uses entire student score histories, including other subjects and teachers, to detect higher than expected student scores associated with particular teachers
Aliases, Variants, Close Extensions	Growth Relative to Self, Raw Gain, Simple Gain, Slope, Average Gain, Gains/Slopes-as-Outcomes, Trajectory Model	Growth-to-Standards Model, Gain-Score Model	Transition Model, Transition Matrix Model, Value Table	Residual Difference Model, Covariate Adjustment Model, Regression Model, Percentile Rank of Residuals	Regression Model, Prediction Model	The Colorado Model, Percentile Growth Trajectories, Conditional Status Percentile Ranks	Sanders Model, EVAAS, TVAAS, Tennessee Model, Layered Model, Variable Persistence Model, Cross-Classified Model
Primary Question(s) 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 future standard?	Is this teacher associated with higher scores for his or her students than expected given all available scores and other teacher effects?
Q1: Primary Interpretation	Growth Description	Growth Prediction	Growth Description and Growth Prediction	Growth Description	Growth Prediction	Growth Description and Growth Prediction	Value Added
Q2: Statistical Foundation	Gain-Based	Gain-Based	Gain-Based	Conditional Status	Conditional Status	Conditional Status	Multivariate
Q3: 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 regression must be met.	Interpretable future scale or future standard.	Large sample sizes for reliable estimation.	For high-stakes value-added uses, many years of student data required for stable teacher effects.
Q4: Group-Level Interpretations	Average gain	Average trajectory or percentage of on-track students	Average across value tables or percentage 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"
Q5: Setting Standards	Requires judgment about adequate gain or adequate average gain. Requires understanding of the scale or can be norm-referenced.	Set by defining a future standard and a time horizon to meet the standard.	Set by defining cut scores for categories and values in value table. Requires judgmental cut scores to define adequacy of both individual and aggregate values.	Requires judgment about adequate residual gain. Requires understanding of the scale or can be norm-referenced.	Set by defining a future standard and a time horizon to meet the standard.	Requires judgment about an adequate SGP or median/average SGP. Predictions require a future standard and a time horizon to meet the standard.	Standards required to support absolute or relative distinctions among teacher/school effects, e.g., awards/sanctions to top/bottom 5%.
Q6: Misinterpretations and Unintended Consequences	Intuitive but dependent on vertical scales that can impart undesired dependencies between growth and initial status or socioeconomic status. Can be inflated by dropping initial scores.	Less of an empirical prediction than an aspirational and descriptive prediction. Requires defensible vertical scale over many years. Can be inflated by dropping initial scores.	Loss of information due to categorization of scores. Requires careful articulation of cut scores across grades and years: assumes an implicit vertical scale. Can be inflated by dropping initial scores.	Not a "gain" but a difference from actual and expected status. Violations of linear regression assumption can lead to distortions. Can be inflated by dropping initial scores.	The "projection" metaphor can be confused with "trajectory" when it is in fact a prediction. Maximizing predictive accuracy can diminish incentives to address low-scoring students.	Sometimes misinterpreted as the percentile rank of gain scores. Sometimes overinterpreted as supporting value-added inferences. Can be inflated by dropping initial scores.	Naming fallacy: calling a metric "value-added" does not make it so. Can be unreliable. Detached from theories about improving teaching. Can be inflated by dropping initial scores.