Comparing Different Methods for Representing and Interpreting Student Growth

Joseph Stevens

Joseph Nese

University of Oregon

and

Ann Schulte

Arizona State University

Presented at the Hawaii International Conference on Education, Honolulu, HI, January, 2013

Contact Information:

Joseph Stevens, Ph.D.
College of Education
5267 University of Oregon
Eugene, OR 97403
(541) 346-2445
stevensj@uoregon.edu

Presentation available at:
http://www.uoregon.edu/~stevensj/HICE2013.pdf
And on NCAASE web site soon: http://www.ncaase.com/

This work was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R32C110004 awarded to the University of Oregon. The opinions expressed are those of the author and do not necessarily represent views of the Institute or the U.S. Department of Education.

Presentation Purpose

- Describe alternative methods for making normative interpretations of student academic growth:
  - Traditional growth norms
  - Student growth percentiles
  - Multilevel growth model norms
- The alternative methods:
  - Depend on different assumptions
  - Have different data requirements
  - Provide different information about student progress
  - Answer different research and policy questions

http://www.uoregon.edu/~stevensj/HICE2013.pdf

Growth Norms Based on Different Ideas of Growth

- Kinds of growth models (Briggs & Betebenner, 2009):
  - Growth conditional on time is an <u>absolute growth model</u>
  - Growth conditional on prior achievement is a <u>relative growth model</u>
- Two methods presented here are examples of absolute growth models:
  - Traditional "medical" growth norms
  - Multilevel model growth norms
- Third method presented here (Student Growth Percentiles) is described as:
  - a relative growth model by Betebenner, 2009
  - a conditional status model by Castellano & Ho, 2012

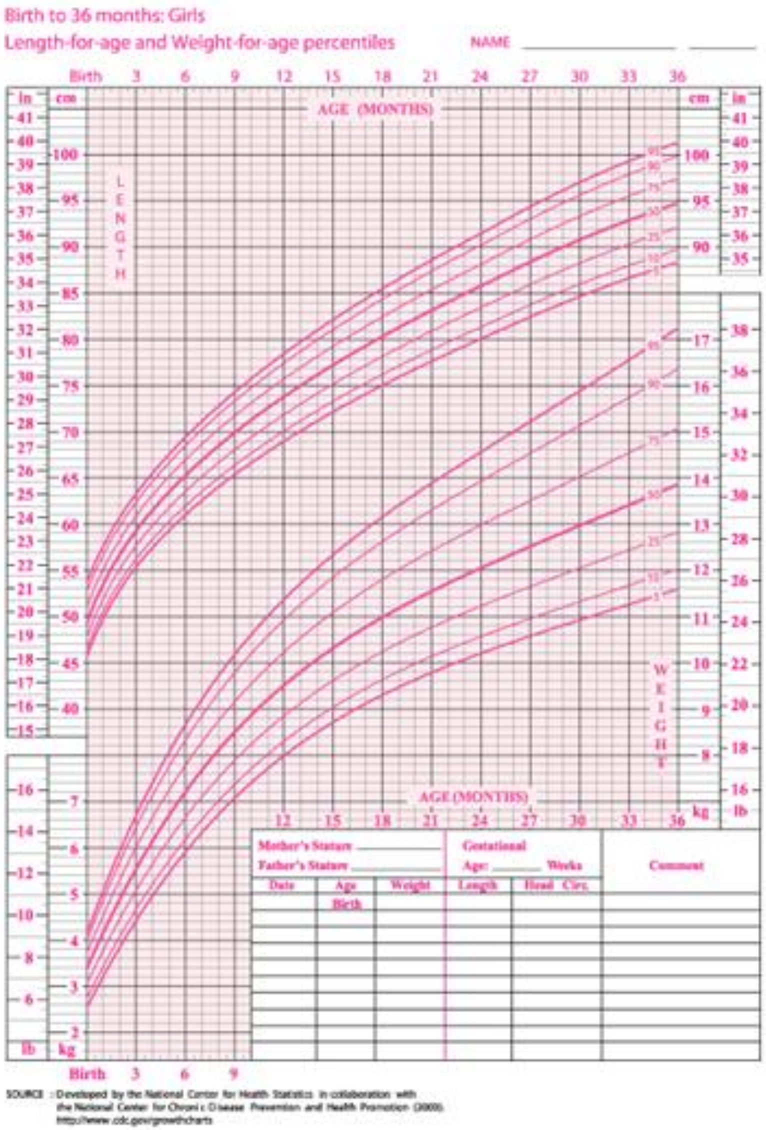
Empirical Examples Presented Here

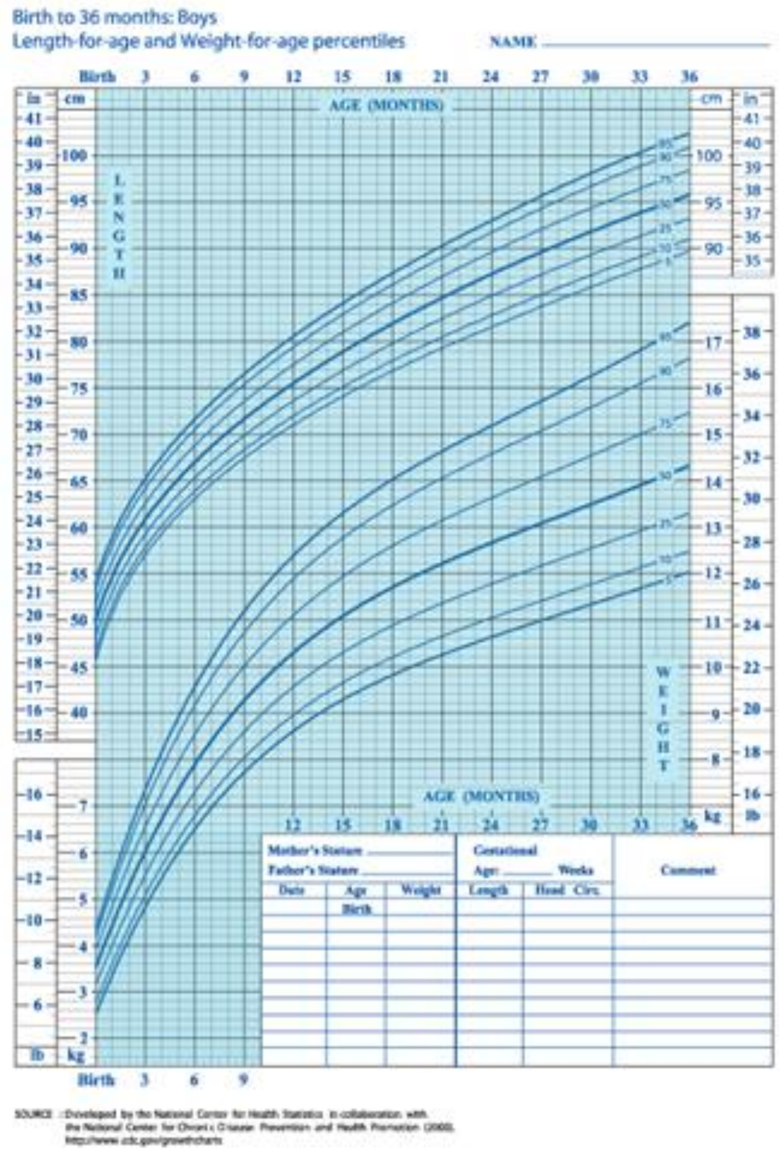
- Based on Oregon state reading/language test scores from a cohort of students who were in the third grade in 2008, 4th grade in 2009, 5th grade in 2010 and 6th grade in 2011
- The complete sample of all students with a valid reading/ language score in 2011 (N = 40,160) had the following characteristics:
  - 49% female
  - 13% current or former LEP students
  - 14% special education
  - 52% economically disadvantaged
  - 66% White, 20% Hispanic, 5% Multi-ethnic, 4% Asian, 3% Black/African American, 2% Native American/Alaskan Native

Traditional Approach to Growth Norms

- “Pediatrician norms”
- Almost always cross-sectional not longitudinal
- Height, weight, stroke risk, fetal growth, etc.
- Interest often in identifying individuals at extremes of “reference” intervals
- Depends on size and representativeness of sample
- Two step procedure used to first smooth curves (e.g., regression), then transform curves to parametric estimates using the LMS (lambda, mu, sigma) procedure (CDC, 2002)
- Used to compare current measurement of an individual to the normative group to evaluate growth or development
- Usually graphical, descriptive interpretation

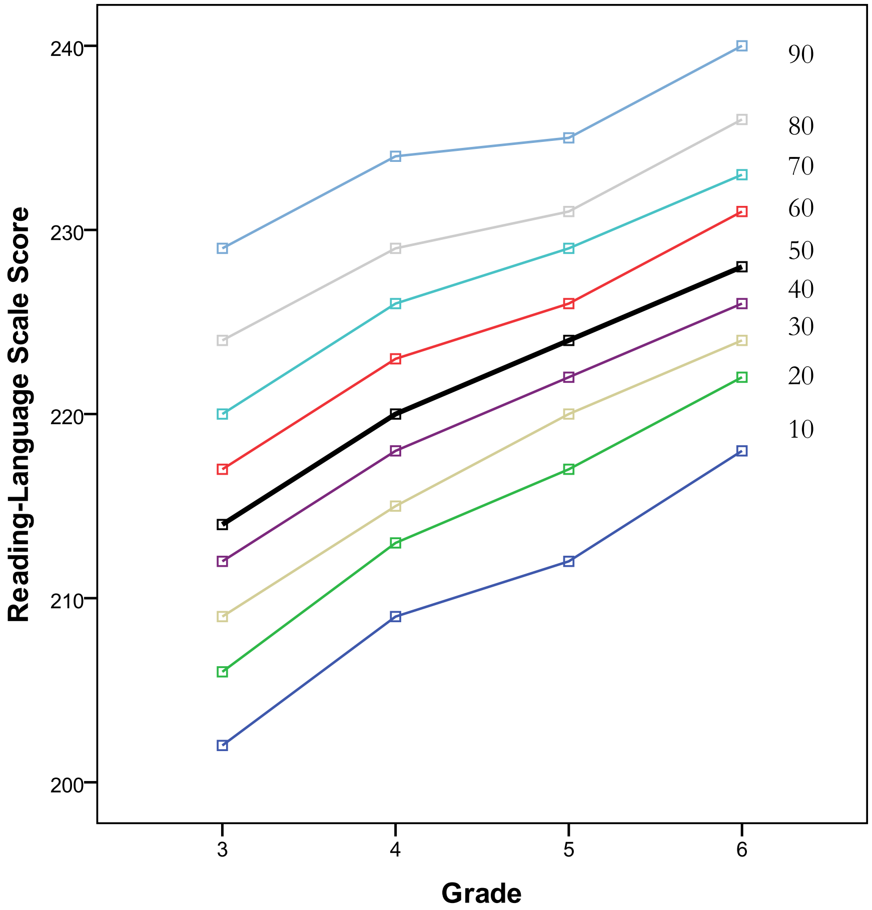
CDC Infant Growth Norms—Length and Weight by Age



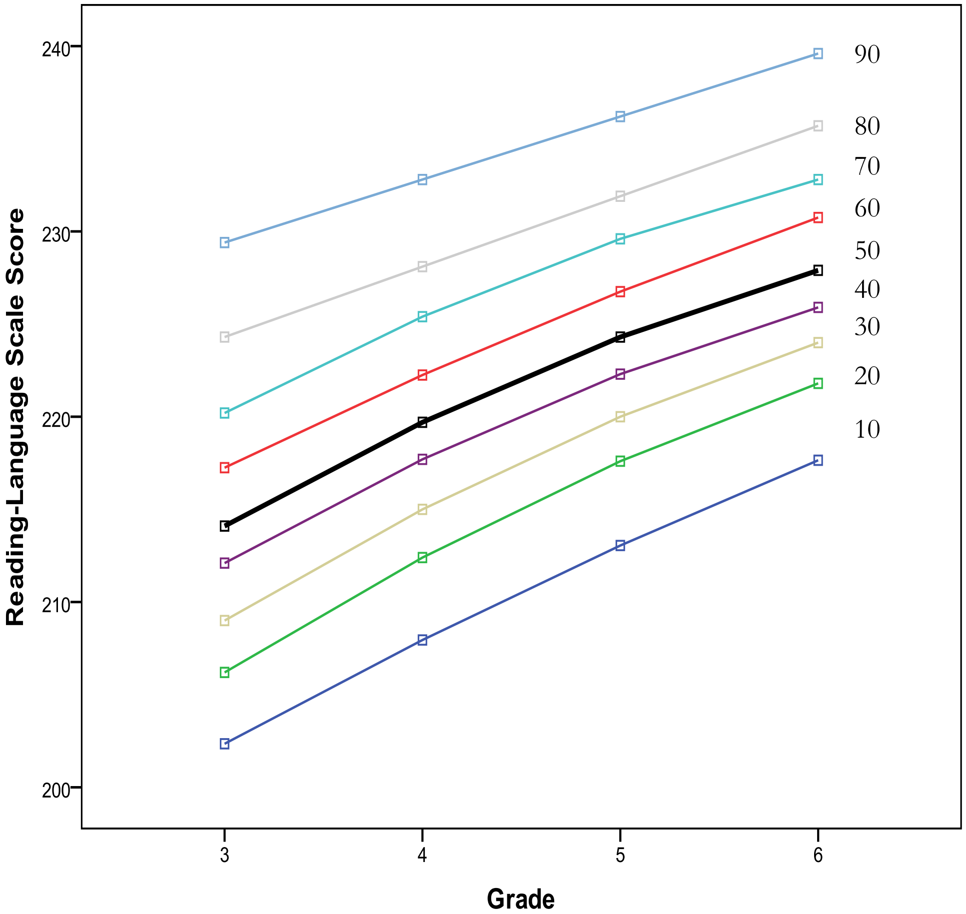


Traditional Approach to Growth Norms

- Following figures present growth norms in deciles (percentile ranks of 10, 20, 30, 40, 50, 60, 70, 80, and 90) for Oregon state test scores
  - Calculated from distribution of reading-language scale scores for all students who had a valid reading-language test score in each year (third grade in 2008, 4th grade in 2009, 5th grade in 2010 and 6th grade in 2011)
  - Cross-sectional sample
- On next slide figure on left shows observed deciles
- Figure on right shows deciles smoothed by regression fitting and a Box-Cox transformation

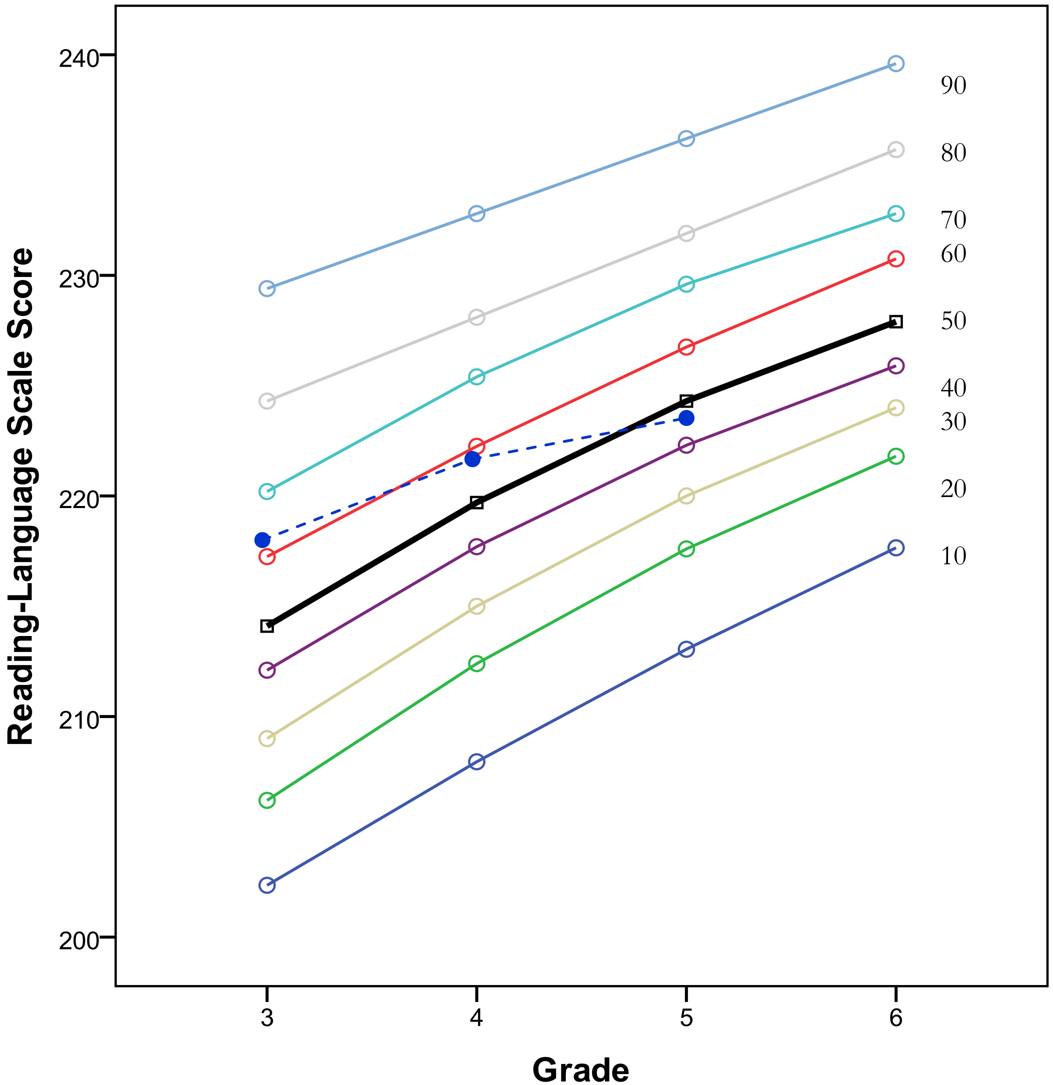


Reading-Language Scale Score
240
230
220
210
200
90
80
70
60
50
40
30
20
10
3
4
5
6
Grade



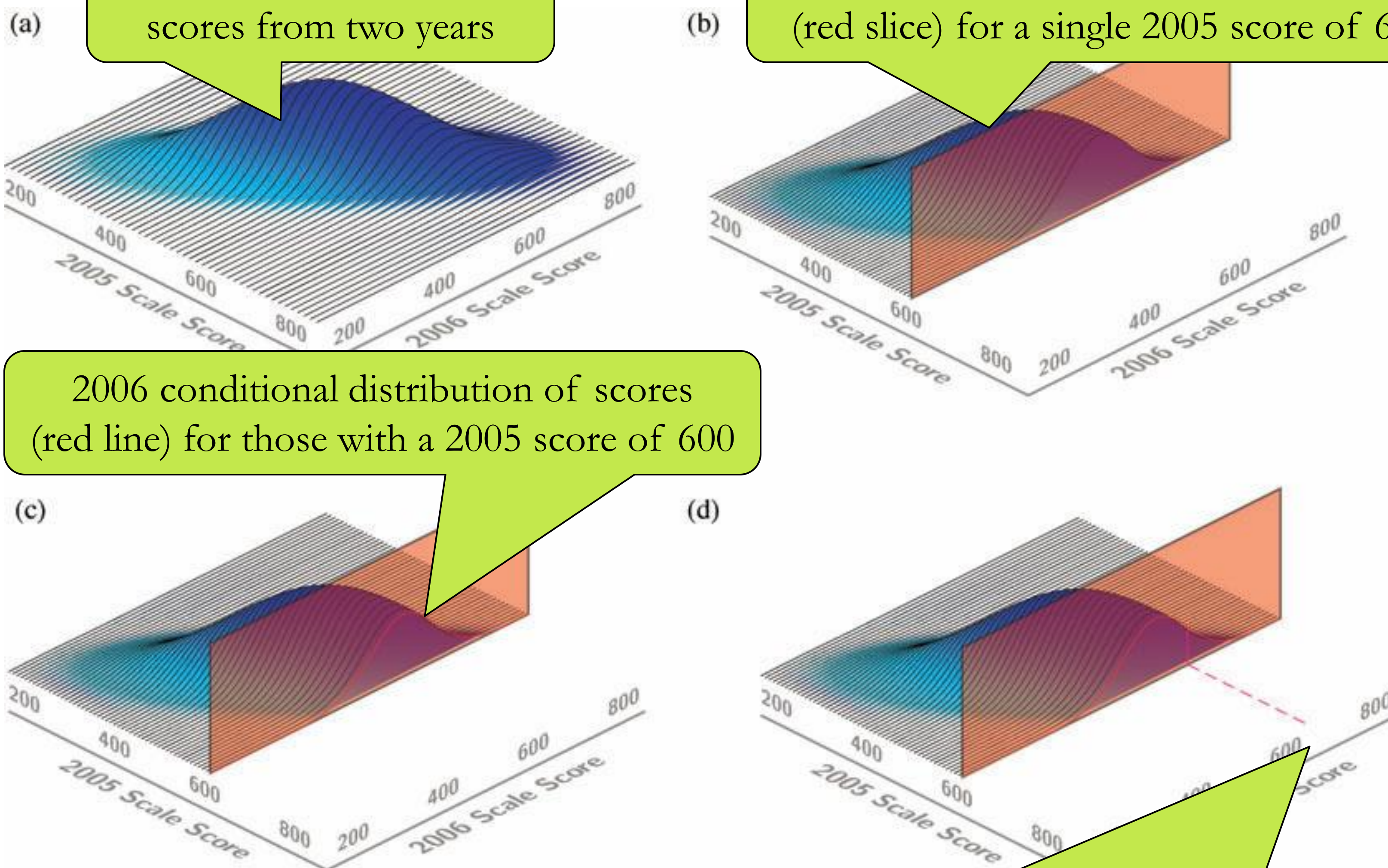
Reading-Language Scale Score
240
230
220
210
200
90
80
70
60
50
40
30
20
10
3
4
5
6
Grade

Example using empirical deciles to interpret an individual student growth curve (dashed line)



Student Growth Percentiles (SGP)

- Described as a Relative Growth Model
  - Current year performance conditioned on prior year(s) of performance
  - Relative rank in a distribution of those who had similar scores in previous years
- Oregon sample composed of all those who had a reading-language score in 2011 and at least one prior year score in years 2008-2010
- Betebenner (2009) approach uses ordinal models (quantile regression) as well as B-spline, cubic polynomial smoothing
- SGP package in R, PROC QUANTREG in SAS



Bivariate distribution of scores from two years

Taking account of prior achievement scores (red slice) for a single 2005 score of 600

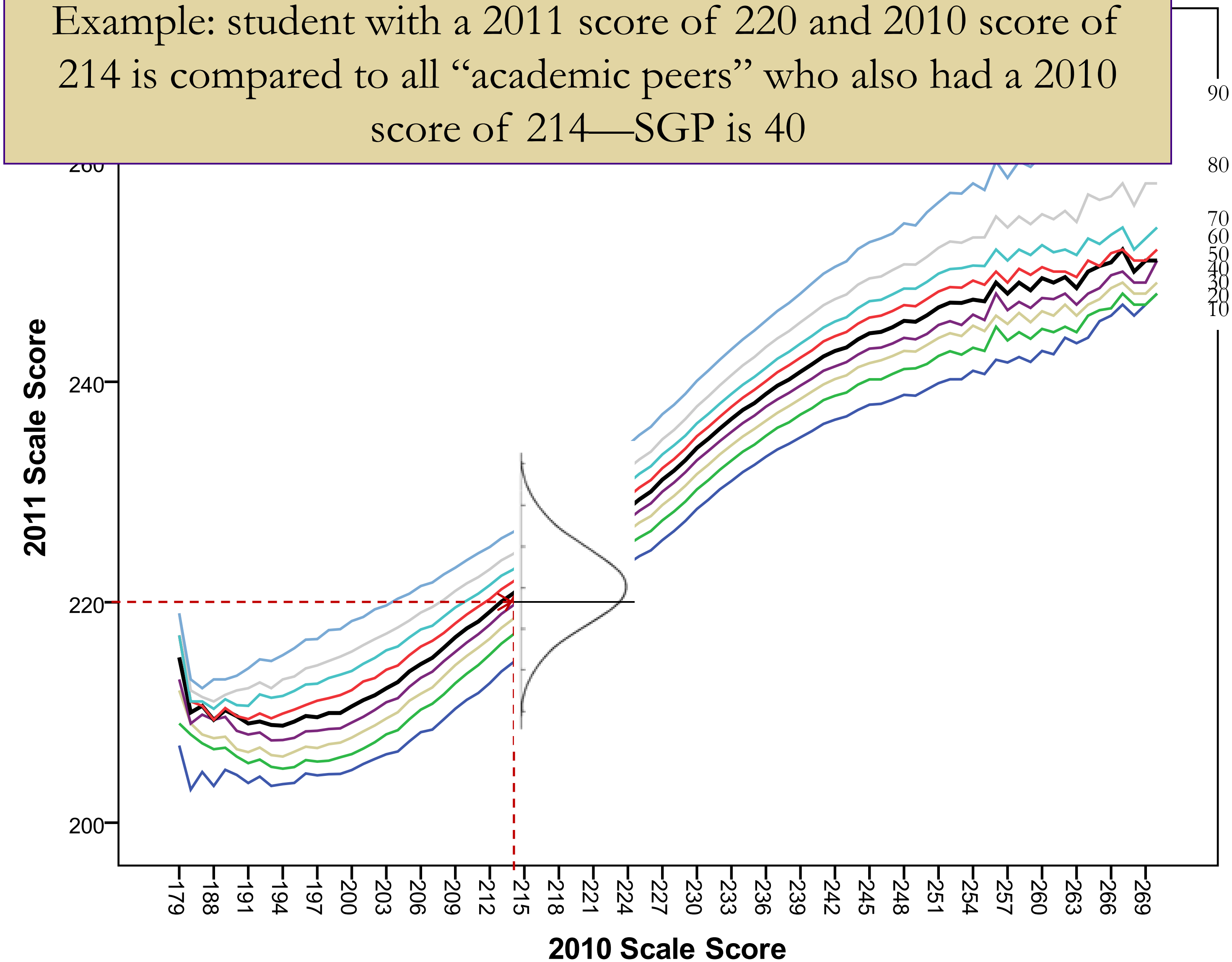
2006 conditional distribution of scores (red line) for those with a 2005 score of 600

For example, a 2006 score of 650 (red dotted line) represents 70th PR for those who had a score of 600 in 2005

Reproduced from Betebenner (

Student Growth Percentiles, Oregon Sample

Example: student with a 2011 score of 220 and 2010 score of 214 is compared to all "academic peers" who also had a 2010 score of 214—SGP is 40



Multilevel Growth Model (MGM) Norms

- Another alternative representation of student growth rests on the statistical modeling of change over time
- These models are absolute growth models in that they relate change to a time function and maintain the metric of the score scale
- Therefore a vertically linked score scale is necessary
- Two types of MGM illustrated here:
  - Two level MGM (time nested within student) with OLS estimation
  - Latent Variable Regression (LVR) in which a latent estimate of intercept is used to predict growth using empirical Bayes estimation

Multilevel Growth Models

Within-person, level-1 (measurement occasions, 1-t):

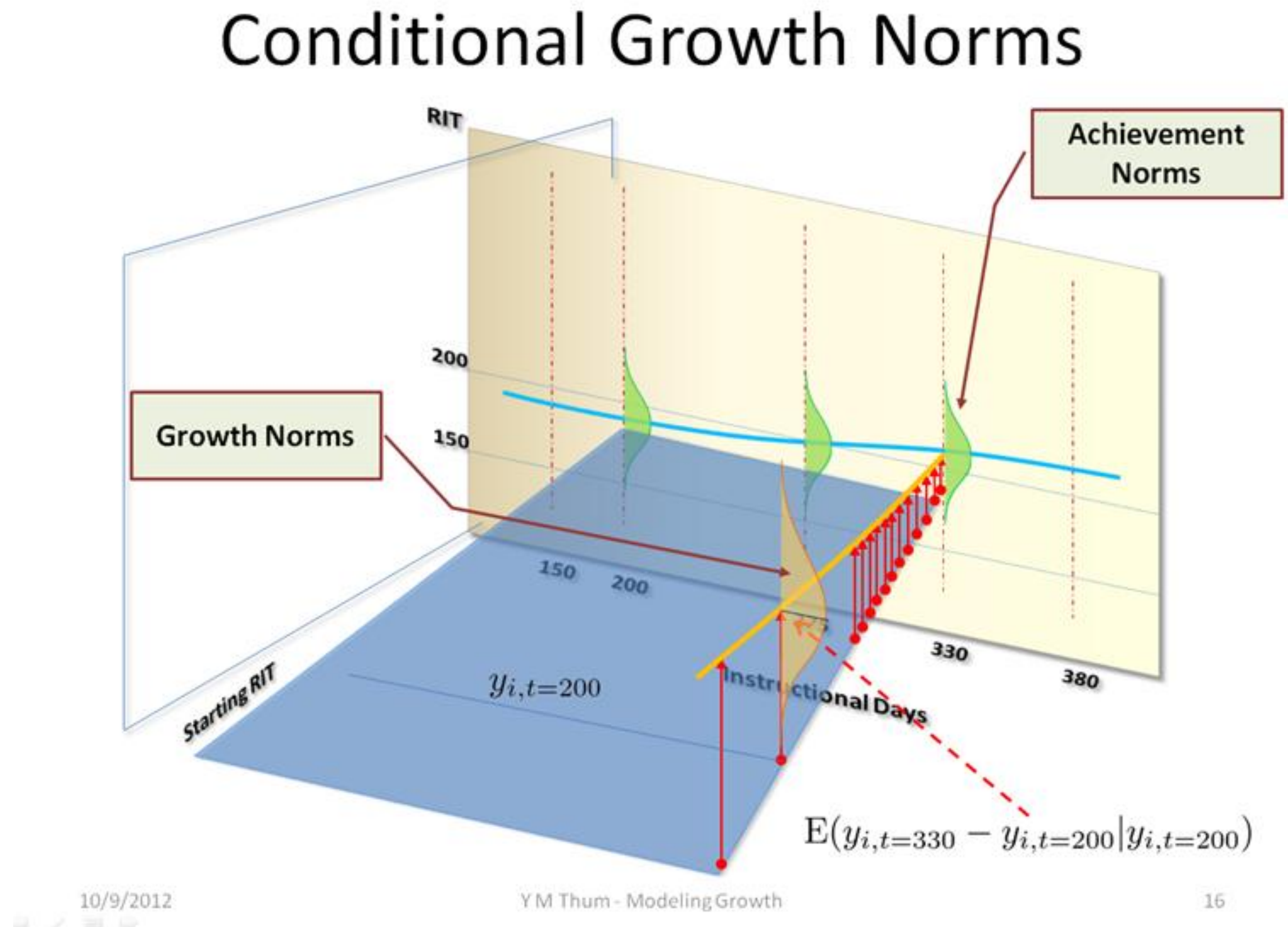
$$\text{Score}_{ij} = \beta_{0j} + \beta_{1j}\,(\text{Time}_{ij}) + \beta_{2j}\,(\text{Time Squared}_{ij}) + r_{ij}$$

Between-person, level-2 (persons, 1-i):

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + u_{2j}$$

Latent Variable Regression (LVR):

$$\beta_{1} = \gamma_{10}^{*} + \gamma_{11}(\beta_{0}) + u_{1}^{*}$$
$$\beta_{2} = \gamma_{20}^{*} + \gamma_{21}(\beta_{0}) + u_{2}^{*}$$



Reproduced from Thum (October, 2012). *The Effective Use of Some School-Level Indicators of Student Learning Growth: NWEA's Learning Productivity Measurement (LPM) System*, 12th Annual Maryland Assessment Conference.

MLM Growth Model Results

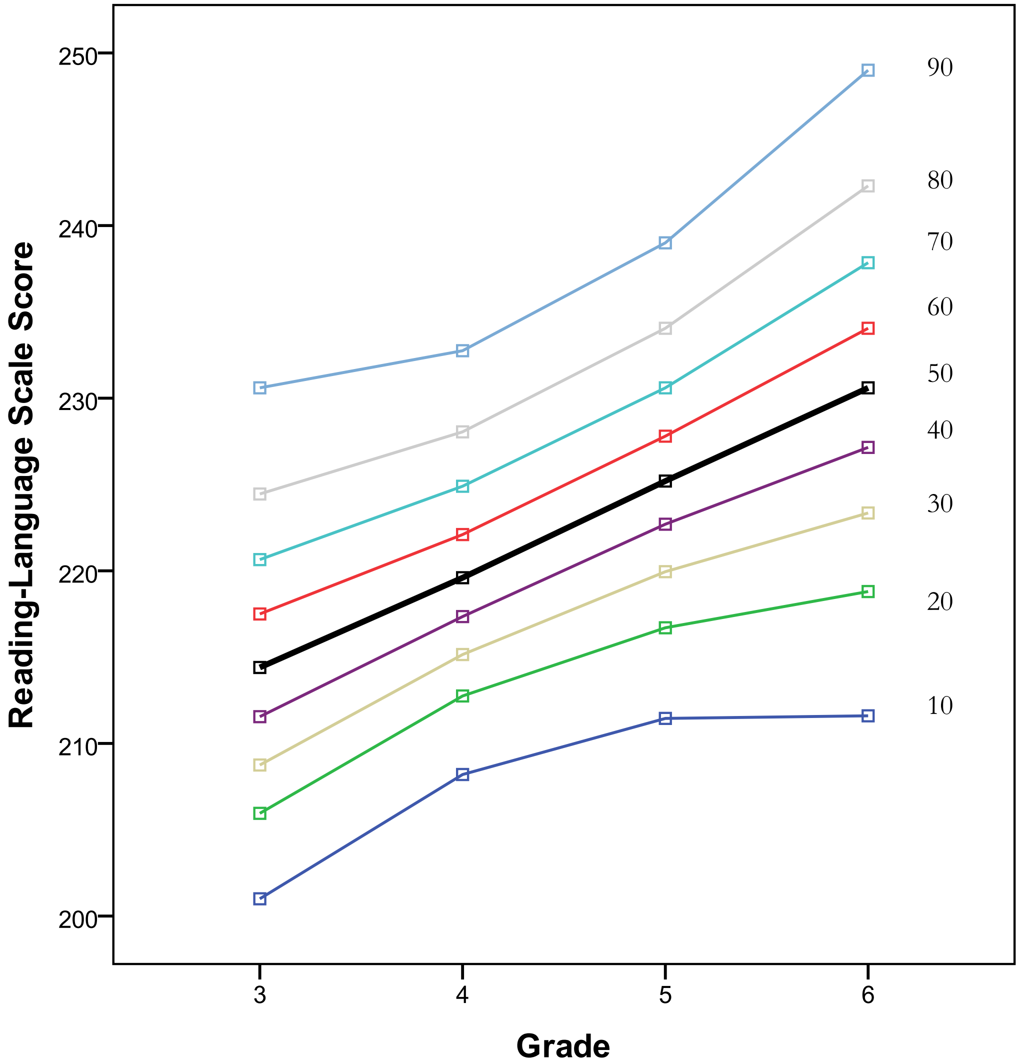
Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	df	p
Intercept, γ_{00}	214.7108	0.0619	3470.395	36948	<0.001
Slope, γ_{10}	5.6303	0.0416	135.381	36948	<0.001
Curvature, γ_{20}	-0.3492	0.0121	-28.951	36948	<0.001

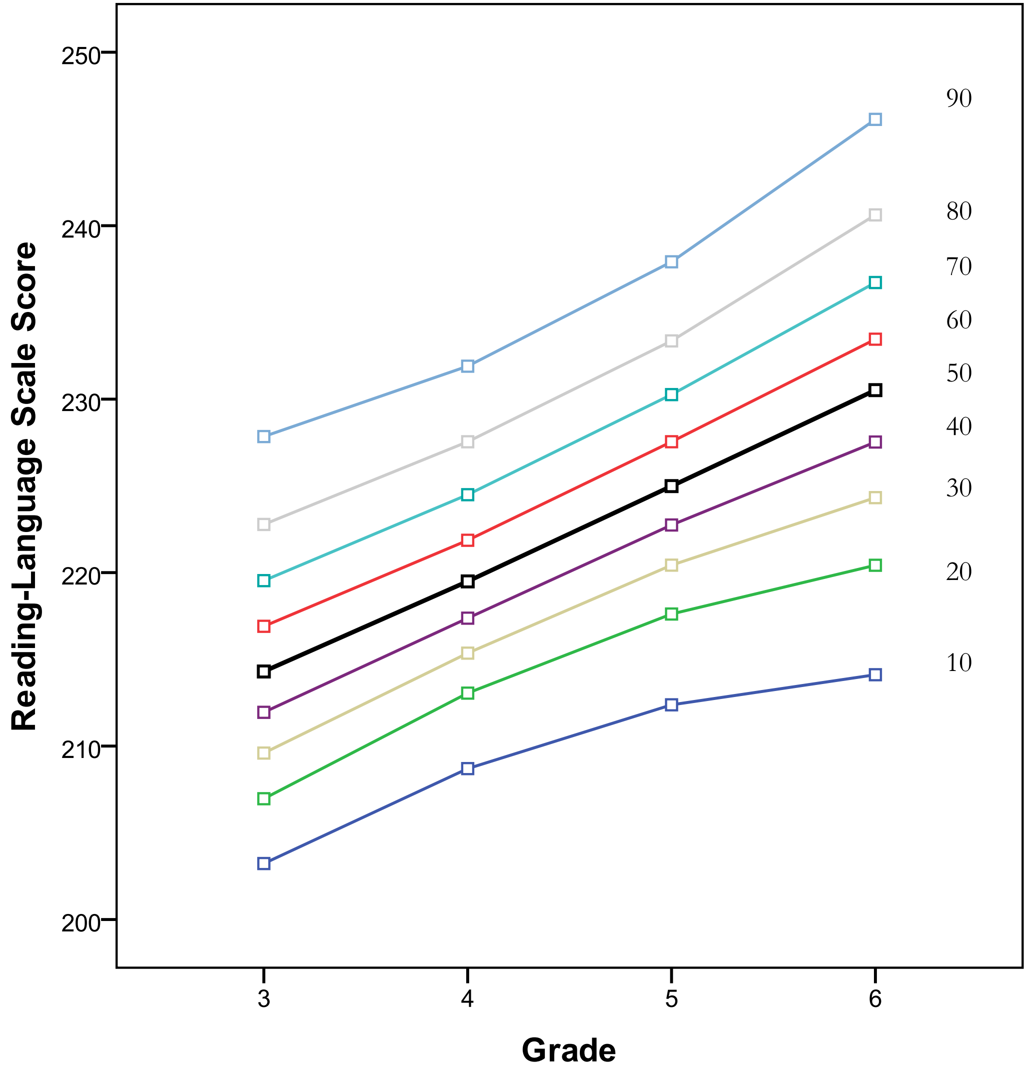
Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	df	χ^2	p-value
Intercept, u_0	10.9042	118.9006	35444	255258.76	<0.001
Slope, u_1	3.8214	14.6029	35444	45773.14	<0.001
Curvature, u_2	0.7920	0.6273	35444	39051.53	<0.001
level-1, r	4.2265	17.8631			

Growth Deciles Based on MLM OLS Estimates



Growth Deciles Based on MLM LVR Empirical Bayes Estimates



Comparing the Methods

- Absolute vs. relative growth methods represent different entities
- What is "growth"?
- Traditional norms:
  - Provide information on absolute growth
  - Based on smoothing and estimation of distribution parameters which are then used to estimate percentiles
  - Assumptions about underlying theoretical distributions lead to use of smoothing methods
  - Explicit evaluation of empirical curves and adaption of methods to ensure fit; sample weighting to ensure representativeness
  - Data requirements: large samples, constant scale over time, cross-sectional
  - Largely descriptive use and interpretation; interpretation straightforward

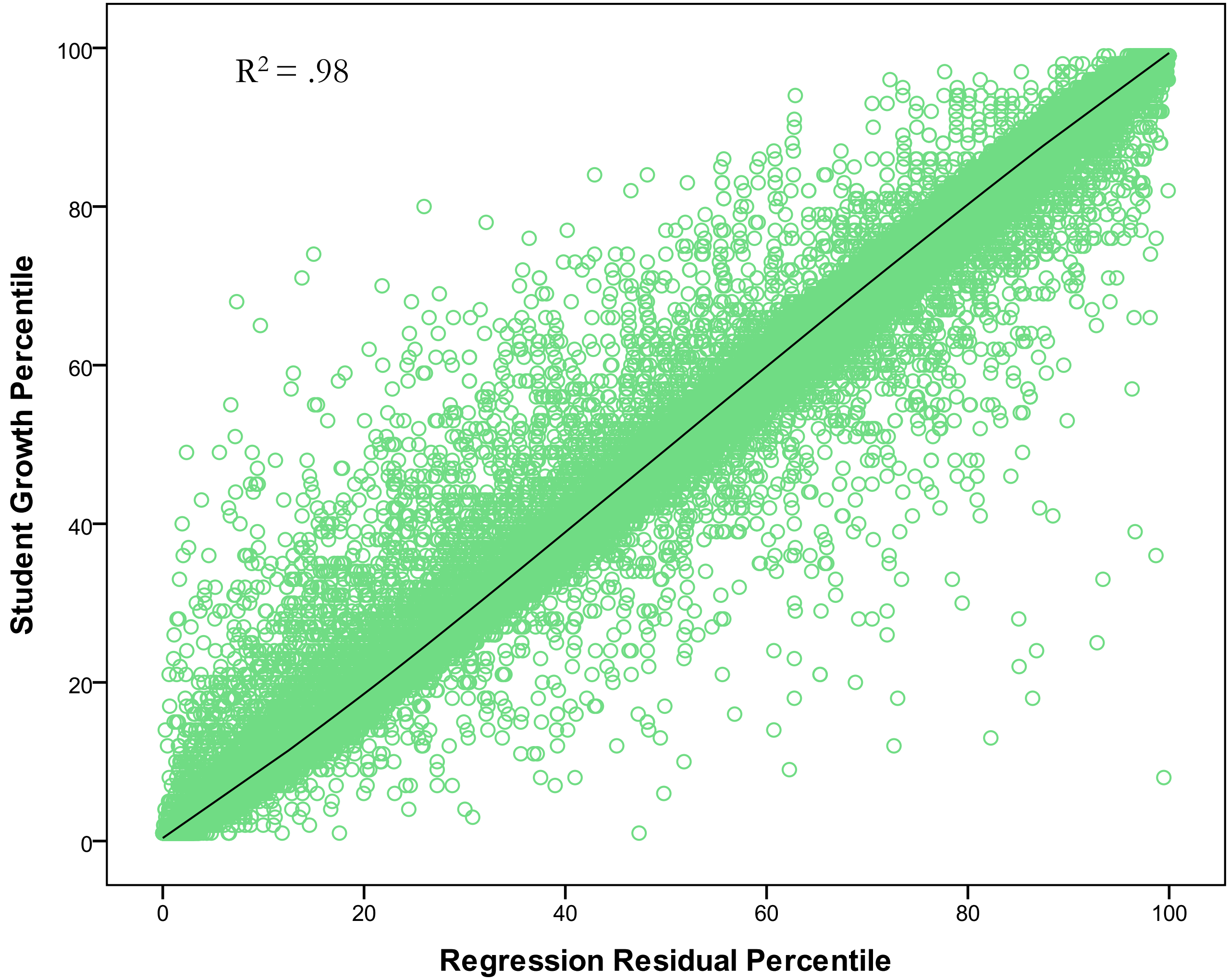
Comparing the Methods

- Student Growth Percentiles:
  - Provide information on relative ranking; do not directly represent growth
  - Based on complex modeling
  - Assume need to correct for scale imperfections and distributional irregularities but same corrections often applied regardless of particular distributional characteristics
  - Data requirements: large samples, do not require same scale (or even content) over time, at least two years of longitudinal data
  - Expression of results in percentile ranks, familiar to users

Comparing the Methods

- Student Growth Percentiles:
  - "Conditional Status Percentile Ranks" a more accurate name than "growth percentiles" (see Castellano & Ho, 2012)
  - Provide only normative information; criterion-referenced interpretations require SGPs to be linked back to score scale or proficiency categories
  - Equivalent to residuals that estimate difference between predicted and actual performance in current year based on previous year(s) test scores
  - Difference mainly in estimation methods
    - SGPs assume ordinal scale and nonnormal score distributions
    - Regression residuals assume interval scale and normal score distributions (median difference in PR = 2.2)

Correlation of SGPs with Conditional Regression Residuals



Comparing the Methods

- MLM Methods:
  - Provide information on absolute growth; LVR models provide information on absolute growth and relative growth simultaneously
  - Based on complex modeling
  - Methods used (MLE and EB) provide highly efficient estimation
  - Data requirements: Moderately sized samples ($N > 200$), constant vertically linked scale over time, longitudinal data, more occasions are better
  - An advantage of MLM is the handling of missing data and the ability to correctly model varying times of measurement
  - Adjusts for unit "weakness" through EB shrinkage
  - Interpretation of deciles is straightforward

Comparing the Methods

- Traditional norms have been well validated especially in some areas of application (e.g., medicine)
- Despite rapid growth in popularity and application, SGPs have little published validity evidence
- Some evidence critical, see Castellano & Ho (2012); not "achievement score gains" but change in relative position
- Substantial evidence regarding use of MLM for estimating teacher and school effects, but little use of MLMs to develop norms
- Need for caution and additional research!

Contact Information

Joseph Stevens, Ph.D.
College of Education
5267 University of Oregon
Eugene, OR 97403
(541) 346-2445
stevensj@uoregon.edu

Presentation available at:

http://www.uoregon.edu/~stevensj/HICE2013.pdf