Alternative Methods for Computing Growth Norms

Joseph Stevens

Joseph Nese

and

Gerald Tindal

University of Oregon

Presented at the annual NCME Conference, Chicago, IL, April, 2015

Contact Information:

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

Presentation available on NCAASE web site: http://www.ncaase.com/

This research was funded in part by a Cooperative Service Agreement from the Institute of Education Sciences (IES) establishing the National Center on Assessment and Accountability for Special Education – NCAASE (PR/Award Number R324C110004); the findings and conclusions expressed do not necessarily represent the views or opinions of the U.S. Department of Education.

Presentation Purpose

- Describe alternative methods for making normative interpretations of student academic growth:
  - Traditional Growth Norms (TGN)
  - Student Growth Percentiles (SGP)
  - Multilevel Growth Model norms (MGM)
- The alternative methods:
  - Depend on different assumptions
  - Have different data requirements
  - Provide different information about student progress
  - Answer different research and policy questions

Growth Norms Based on Different Ideas of Growth

- Kinds of growth models (Briggs & Betebenner, 2009):
  - Growth conditional on time is an absolute growth model
  - Growth conditional on prior achievement is a relative growth model
- Two methods presented here are examples of absolute growth models:
  - Traditional Growth Norms (TGN)
  - Multilevel Growth Model (MGM) 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, 2013

Empirical Examples Presented Here

- Based on three cohorts of fifth grade students in a large school district in Arizona
- 3,985 students across the three cohorts; 3,949 students (99%) had at least one mathematics score and 3,947 students (99%) had at least one reading score
- Sample demographics:
  - White (23%), Hispanic (52%), Black/African American (11%), and American Indian/Alaskan Native students (8%); majority-minority district
  - 75% eligible for FRL
  - 36% active or monitored ELL
  - 48% female

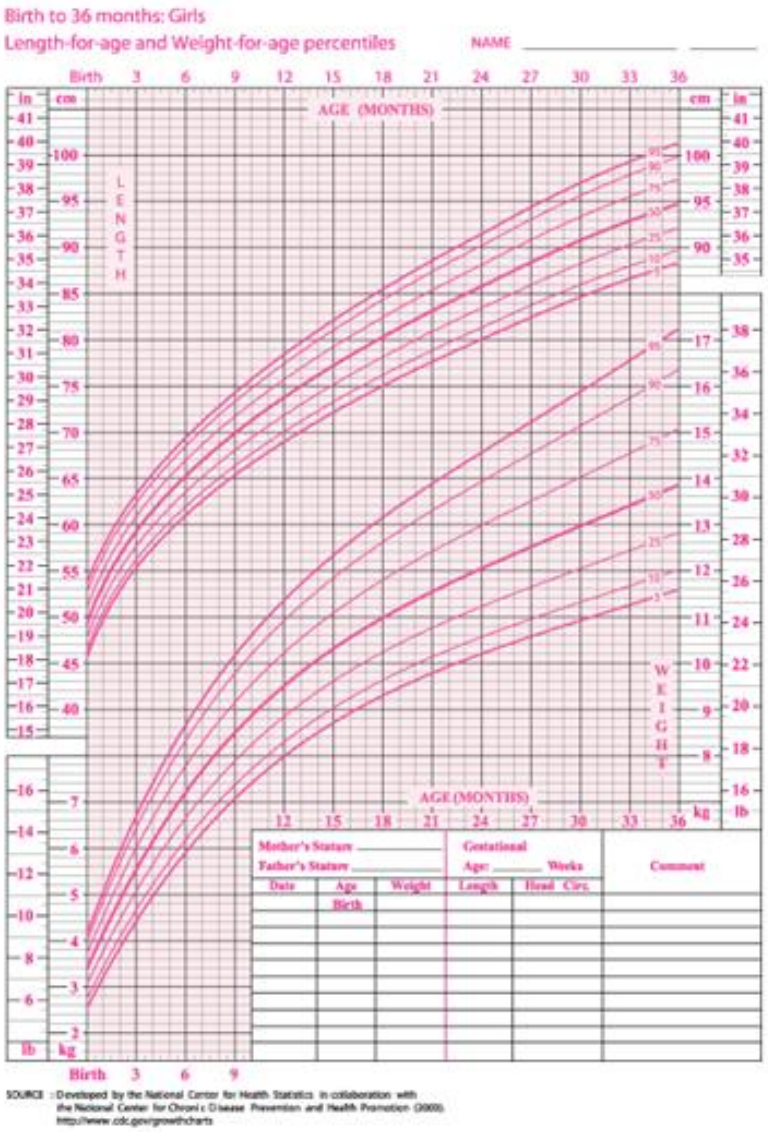
Instrument

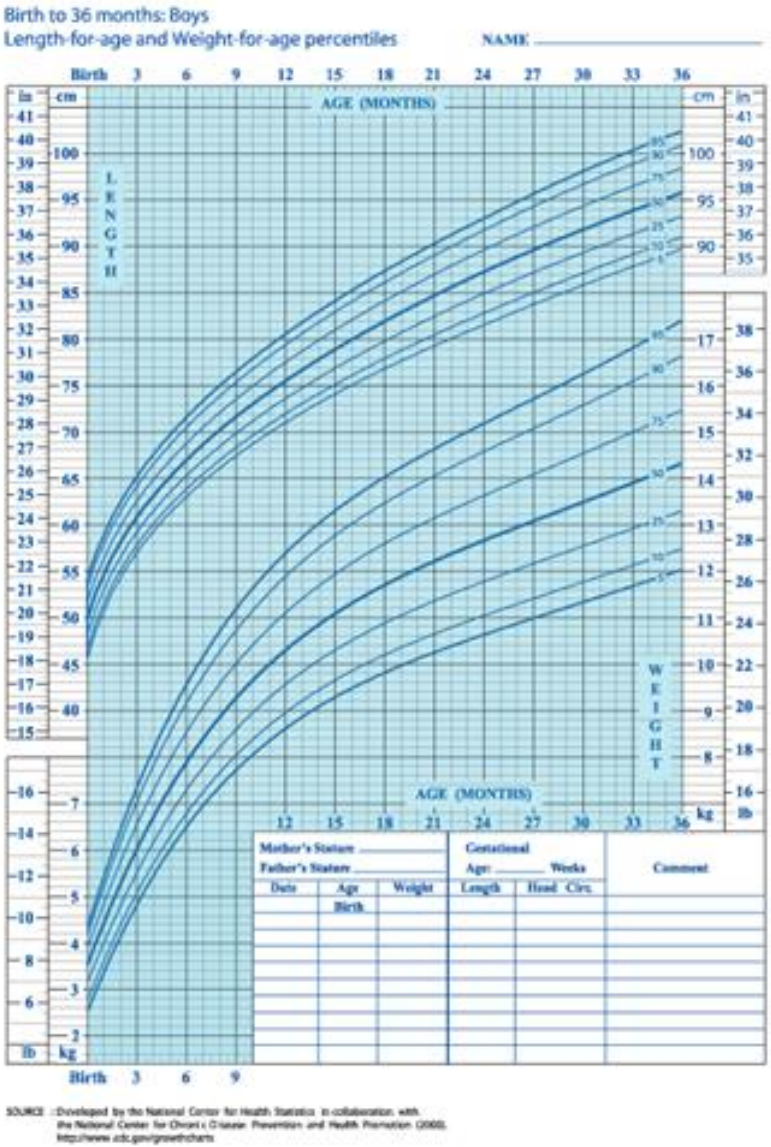
- Student scores on the mathematics and reading subtests of the Measures of Academic Progress (MAP; Northwest Evaluation Association, 2011)
- MAP was administered seasonally (fall, winter, spring)
- MAP is an untimed computerized adaptive test
- Each test includes 50 multiple-choice items with 4 or 5 response options
- MAP items calibrated on a common, vertical scale, using a one parameter, IRT (Rasch) model (NWEA, 2011)

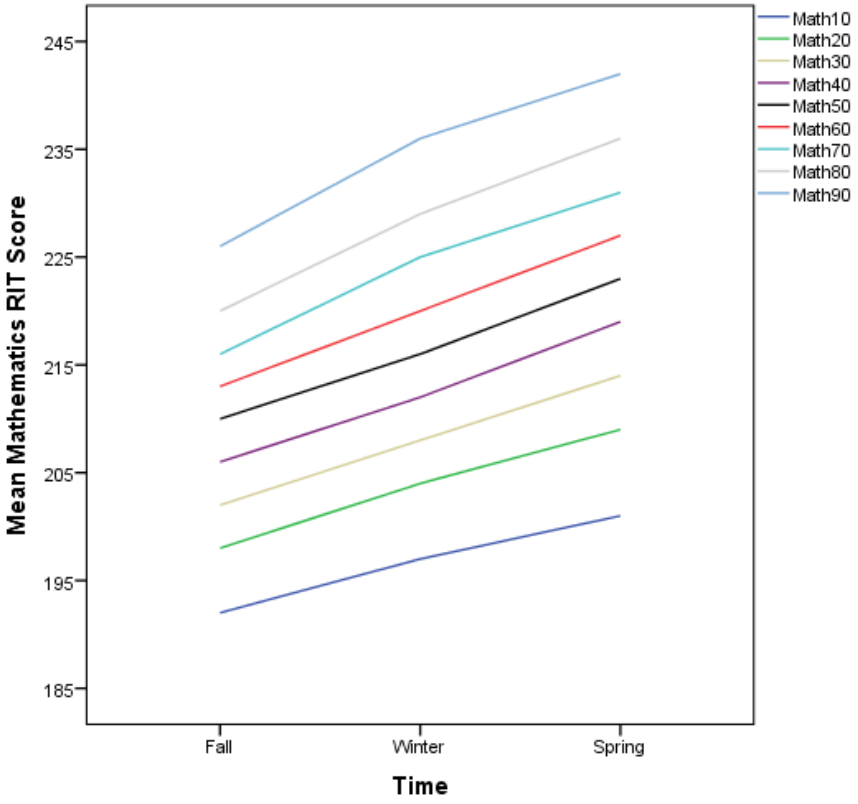
Traditional Approach to Growth Norms

- Used with measures of intelligence, achievement, social competence, psychological functioning, health science applications, etc.
- Interest often in identifying individuals at extremes of "reference" intervals or in tracking developmental progress
- Key consideration is the sampling method used to select the norms group
- Another important consideration is local vs. "population" norms
- Empirical descriptions of norm group performance often augmented by fitting regression models and/or smoothing of distributions (we present unsmoothed deciles here)

CDC Infant Growth Norms—Length and Weight by Age







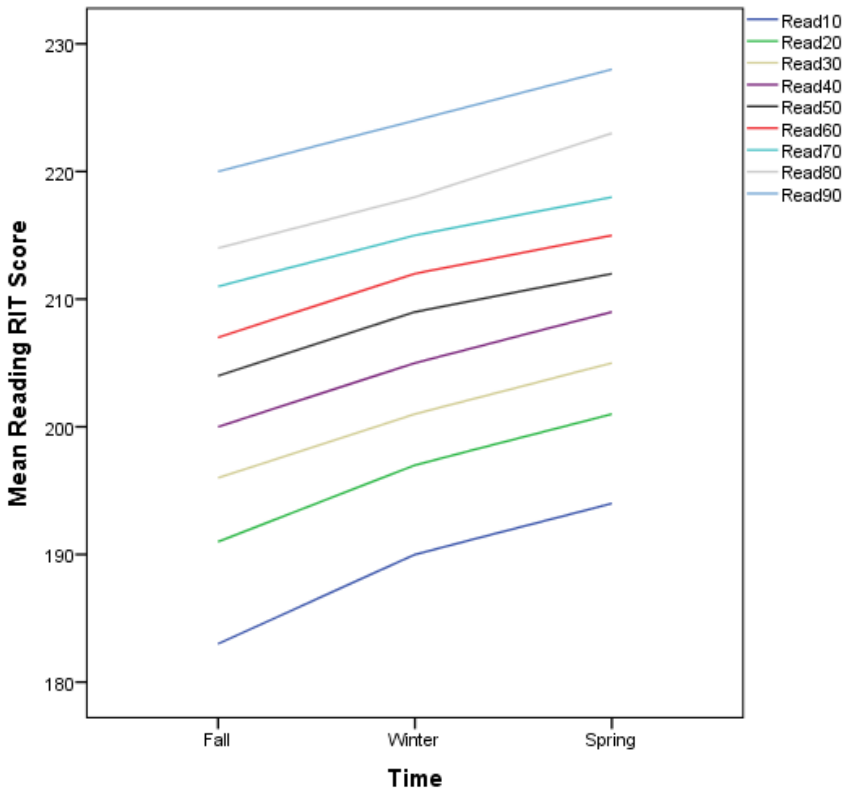
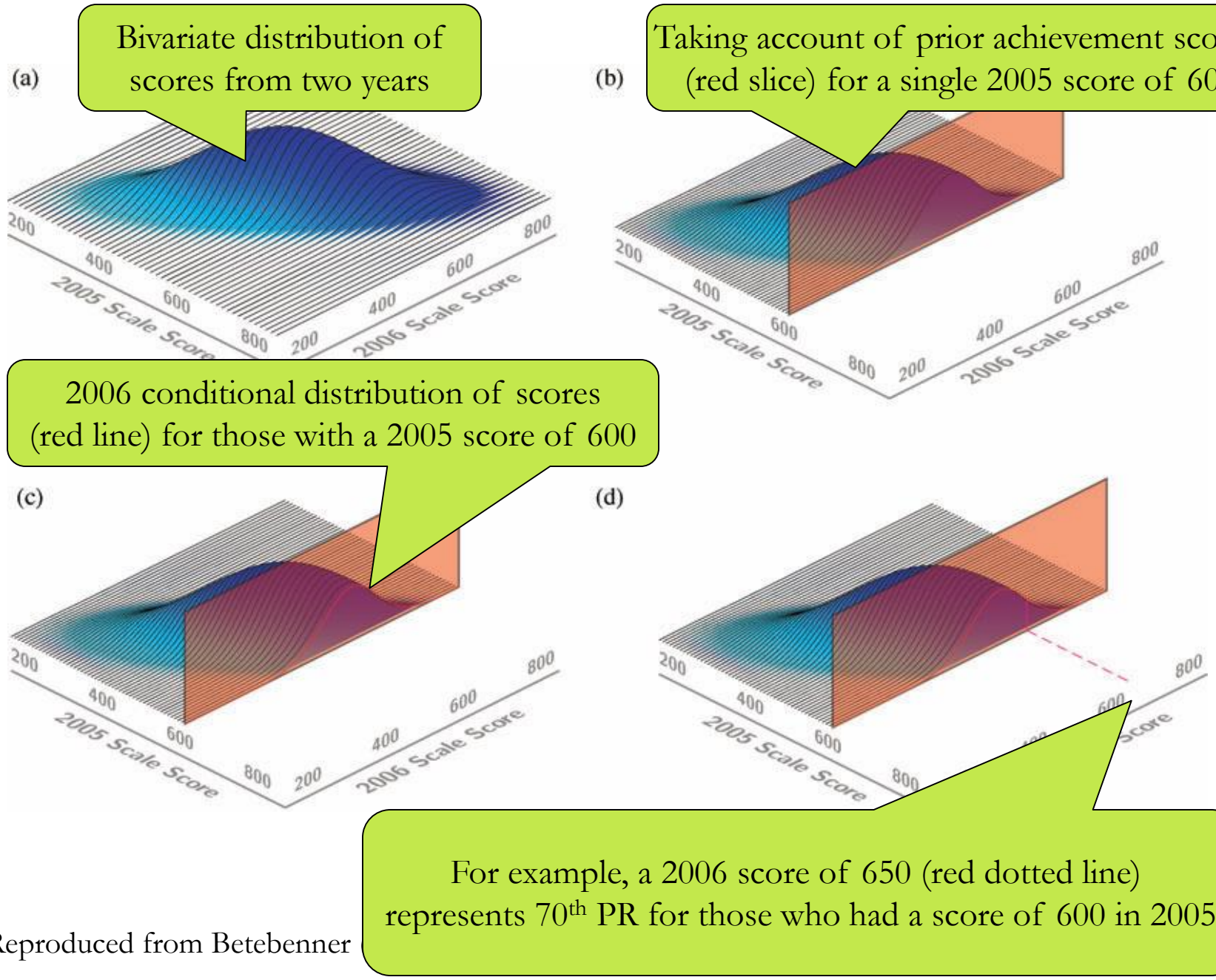


Figure. Traditional growth norm deciles for mathematics (panel on left) and reading (panel on right).

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
- In essence a gain score model; also referred to as a "conditional status percentile rank" model by Castellano & Ho (2013)
- Betebenner (2009) approach uses ordinal models (quantile regression) as well as polynomial smoothing
- We used SGP package in *R*



Reproduced from Betebenner

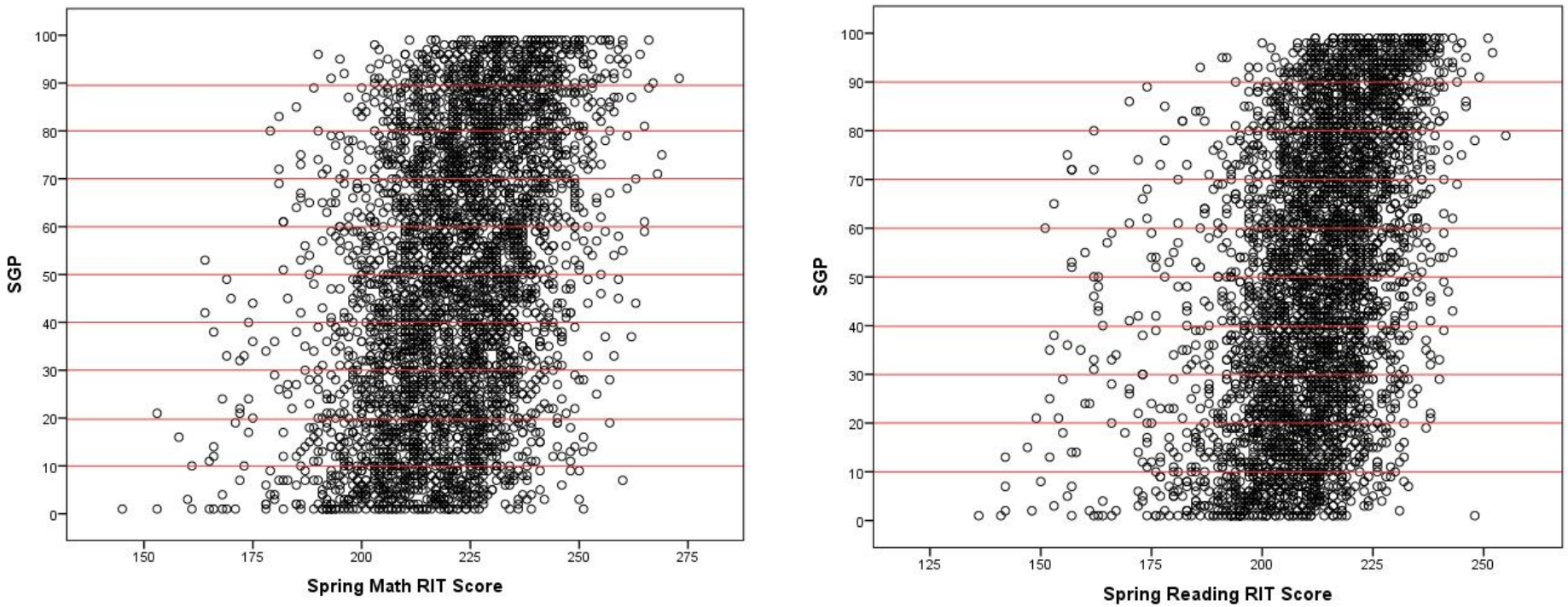


Figure. SGP norm deciles for mathematics (panel on left) and reading (panel on right).

Multilevel Growth Model (MGM) Norms

- Multilevel modeling of change over time
  - Absolute growth model
  - Requires a vertically linked score scale
- Two-level MGM used here using HLM 7
- Time coded 0, 1, or 2 for fall, winter, and spring test administrations

Level-1 (measurement occasions, 1-t):

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

Level-2 (persons, 1-i):

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

- We calculated deciles of both OLS and EB estimates

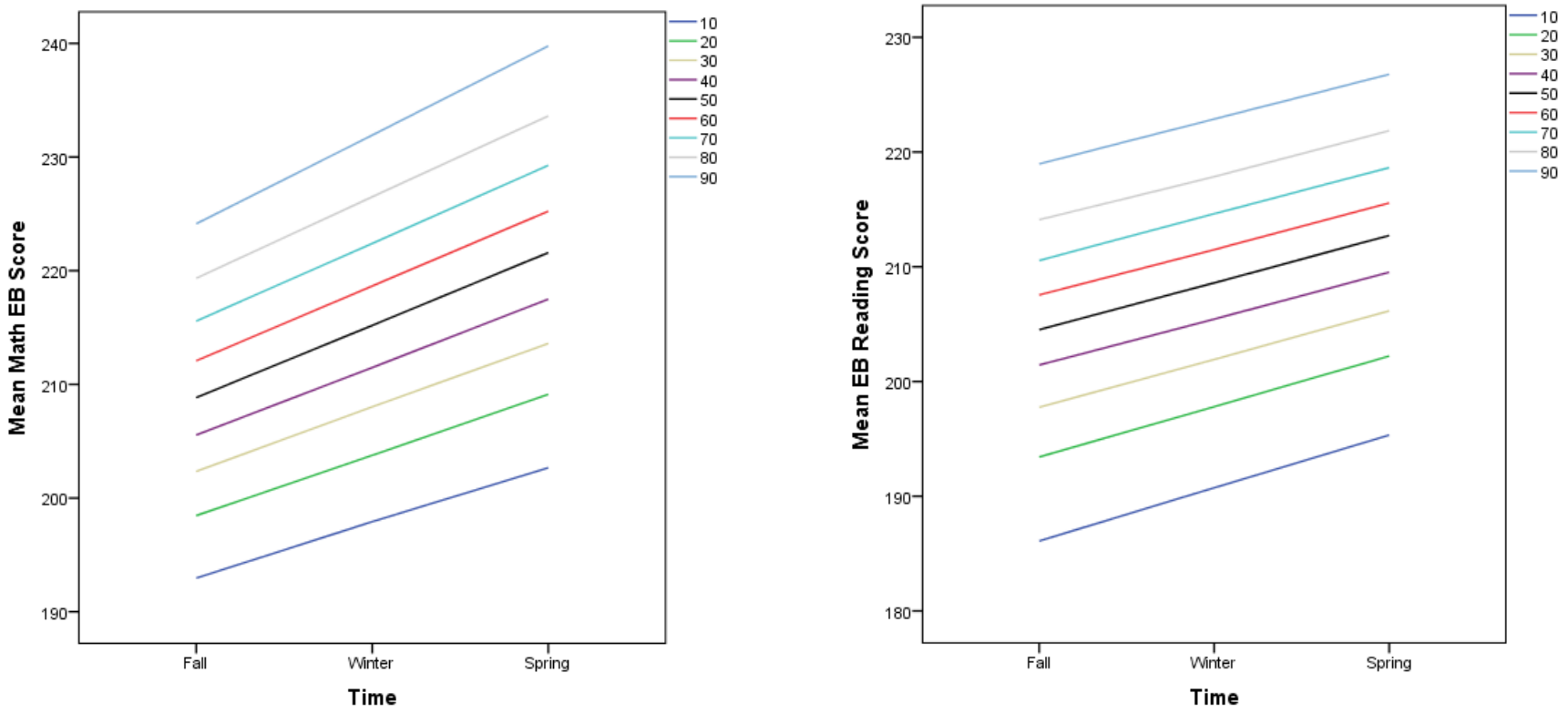


Figure. MGM empirical Bayes deciles for mathematics (panel on left) and reading (panel on right).

Comparing the Methods

- Following Castellano & Ho (2013) we also calculated conditional regression residuals (CSR) to compare to SGP results

- Correlations: Math in below diagonal, Reading above diagonal

	TGN spring	MGM spring	SGP spring	CSR spring	OLS slope	EB slope
TGN spring	1.00	.96	.46	.44	.21	-.09
MGM spring	.98	1.00	.19	.18	-.02	-.31
SGP spring	.39	.20	1.00	.99	.81	.70
CSR spring	.40	.21	.99	1.00	.80	.69
OLS slope	.50	.35	.82	.82	1.00	.93
EB slope	.94	.87	.55	.57	.74	1.00

Comparing the Methods

- Random sample of student PRs by method in math and reading:

Student	TGN	MGM	SGP	CR	OLS slope	EB slope
A	43	55	10	13	8	25
B	34	37	34	36	29	30
C	73	60	63	65	43	60
D	66	39	79	78	33	47
E	73	84	21	22	38	64

Student	TGN	MGM	SGP	CR	OLS slope	EB slope
F	57	79	2	5	9	6
G	29	40	11	12	18	19
H	45	50	32	33	30	29
I	94	96	49	46	60	40
J	51	54	42	40	74	74

Comparing the Methods

- Absolute vs. relative growth methods represent different entities
- What is "growth"?
- Traditional norms:
  - Provide information on absolute growth
  - Assumptions about underlying theoretical distributions often lead to use of smoothing methods
  - Data requirements: representative, large samples; vertical scale
  - Largely descriptive use and interpretation; interpretation straightforward
  - Allows absolute judgment of level of performance at any point in time included in the norms

Comparing the Methods

- Student Growth Percentiles:
  - Provide information on relative ranking; do not directly represent growth; in essence a "residual gain score"
  - Based on complex modeling
  - Assumes 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
  - Misinterpretations likely:
    - The term "growth" is used, but not really growth
    - Documents say SGP provides comparisons to "student's peer group," but not peers in any usual sense

Comparing the Methods

- MGM Methods:
  - Provide information on absolute growth
  - Based on complex modeling; methods used (MLE and EB) provide highly efficient estimation
  - Data requirements: Moderately sized samples ($N > 200$), vertical scale, longitudinal data, more occasions are better
  - An advantage of MGM is the handling of missing data and the ability to correctly model varying times of measurement
  - Adjusts for unit "weakness" through EB shrinkage
  - Allows absolute judgment of level of performance at any point in time included in the norms
  - Interpretation of deciles is straightforward
- Need for caution, careful interpretation, and additional research on all methods!