

The Educational Measurement  
*Statistic De Jour:*  
Student Growth Percentiles  
(yes, that's Comic Sans font...)

# Coming to a State Near You...

## APPENDIX A: WHERE THE COLORADO GROWTH MODEL IS USED

**Table A.1. States Using the Colorado Growth Model in Teacher Evaluations**

State	SGP Measure	Level of Aggregation	First School Year for Evaluations
Colorado	Median	Local education agency decision	2013–2014
Hawaii	Median	Teacher	2014–2015
Massachusetts	Median	Teacher	2013–2014 (some schools starting in 2011–2012)
New Jersey	Median	Teacher	2013–2014
New York	Average	Teacher	2012–2013 (except New York City)
Rhode Island	Median	Teacher	2014–2015
Virginia	Median	Local education agency decision	2012–2013
West Virginia	Median	School	2013–2014

Source: State education agency websites.

- **Add Georgia, Idaho, Nevada, Washington, and, I'm guessing, others.**

*From How Does a Value-Added Model Compare to the Colorado Growth Model?*

- October 2013

From the  
**MDE - AdvancED Michigan**  
**2014 Fall School Improvement Conference**  
November 18, 2014

- Unlike performance level change (PLC) indices, MGP can be easily/validly used across assessment transition years

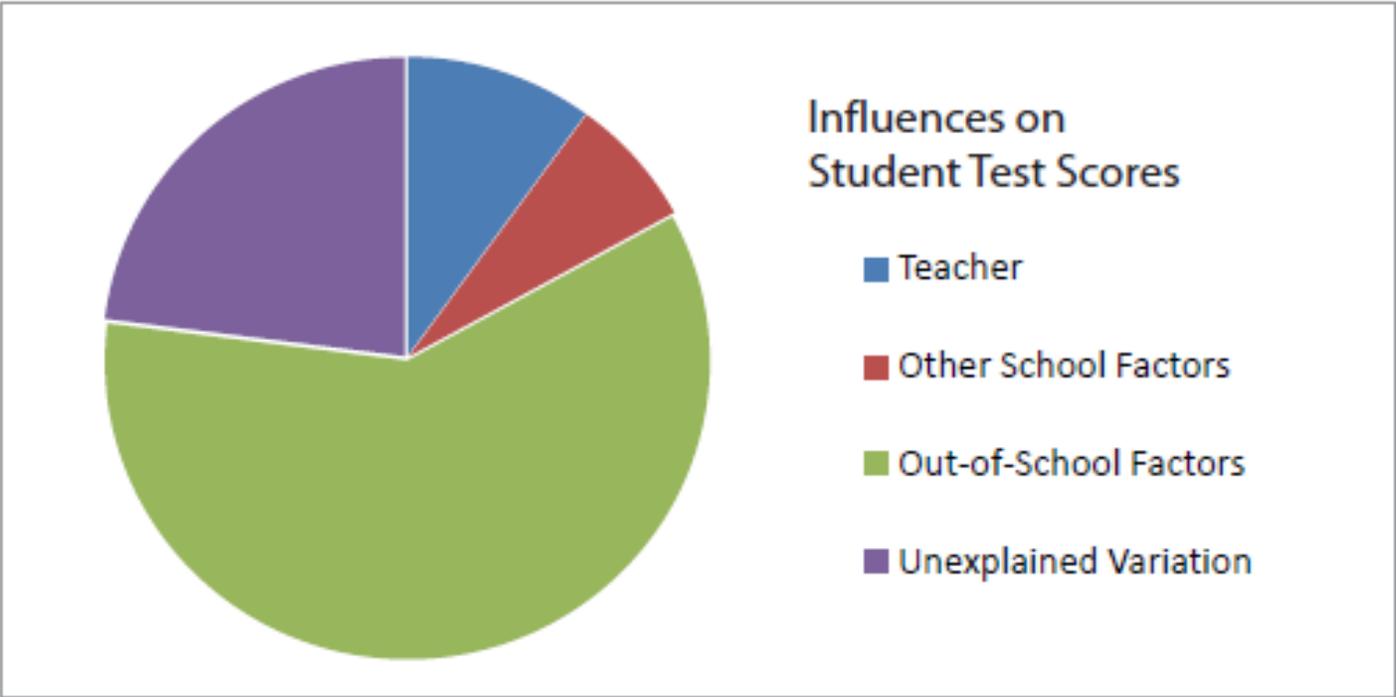
For what?...Instruments (or indices), aren't valid or invalid – it's the purpose to which they're put that determines validity (or lack thereof)

- Like SGPs, MGPs can theoretically range from 0-99, but unlike SGPs, extreme high or low values are unlikely

Not sure the point, but a range of 40-60 can just as easily produce ranks as a range of 0 to 99

A nice, general, pie-chart representing what we know about student test score "growth"

**Figure 1**  
**How Much Variance in Student Test Score Gains Is Due to Variation Among Teachers?**



Reliability and Validity of Inferences About Teachers Based on Student Test Scores, Edward Haertel, ETS's *William H. Angoff Memorial Lecture Series*

# Determining *Average* Growth

## Median Points for Each Scale Score Starting Point

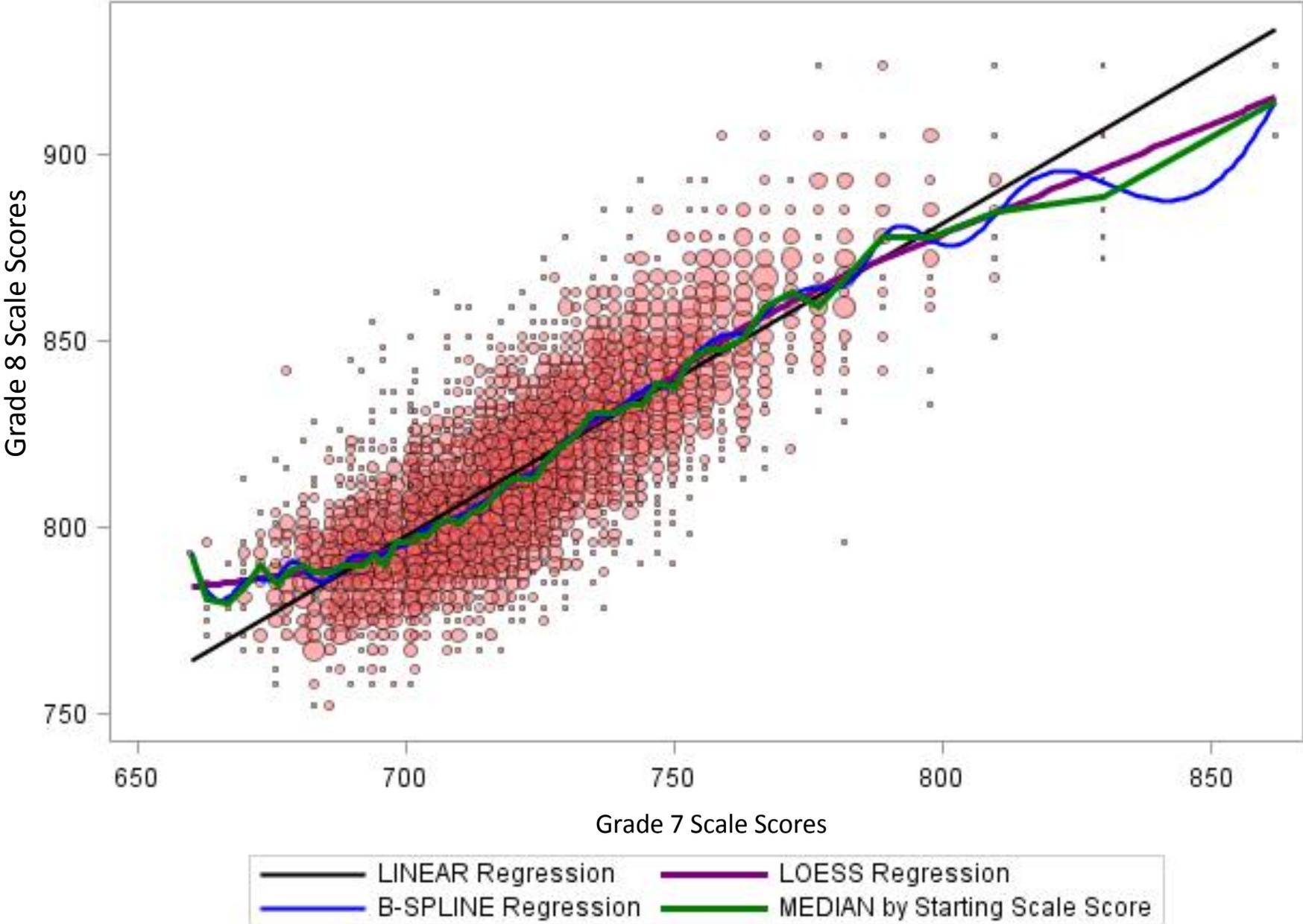
MEAP MATH, Grade 7 to Grade 8



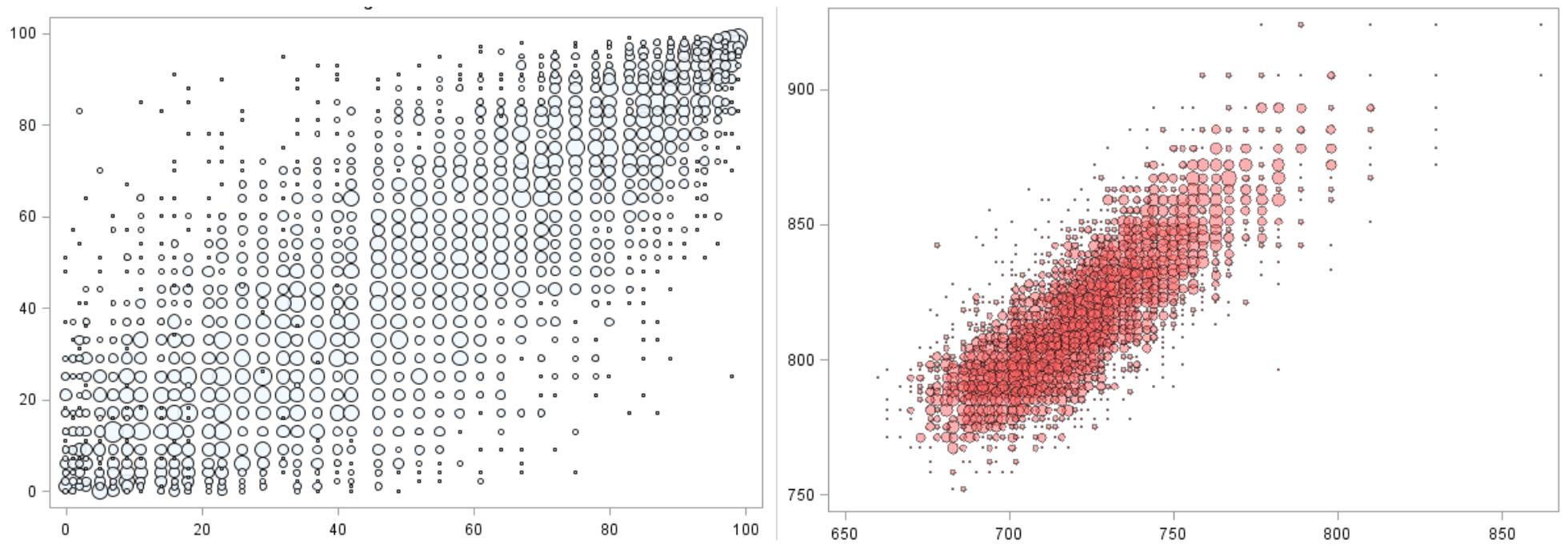
Data: BAA Secure site downloadable Scorecard files -- Schools in Genesee ISD and Lapeer ISD

# Comparing Several Methods for Determining Average Growth

MEAP MATH, Grade 7 to Grade 8

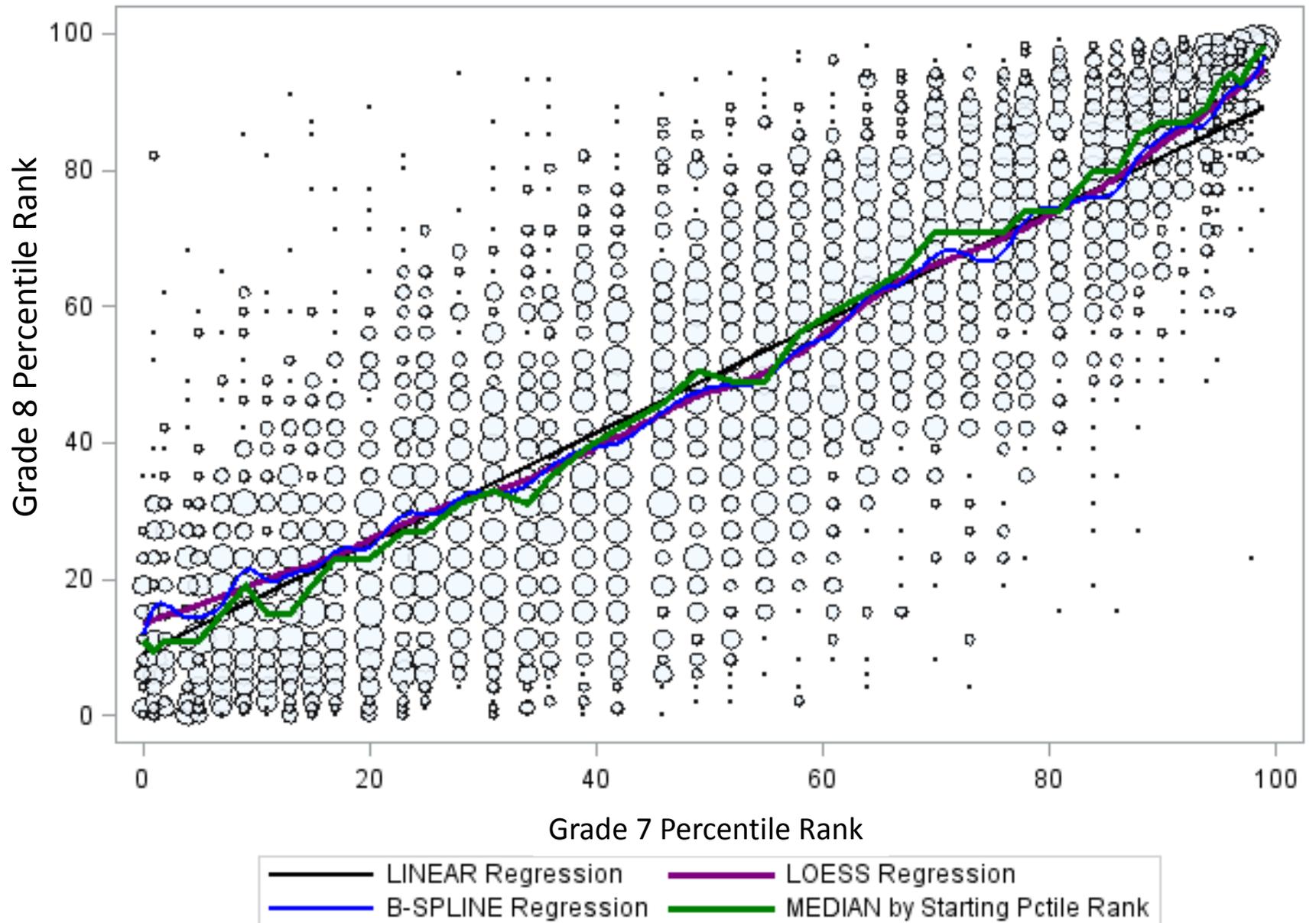


What does the scatterplot look like, when scores are translated to percentiles?



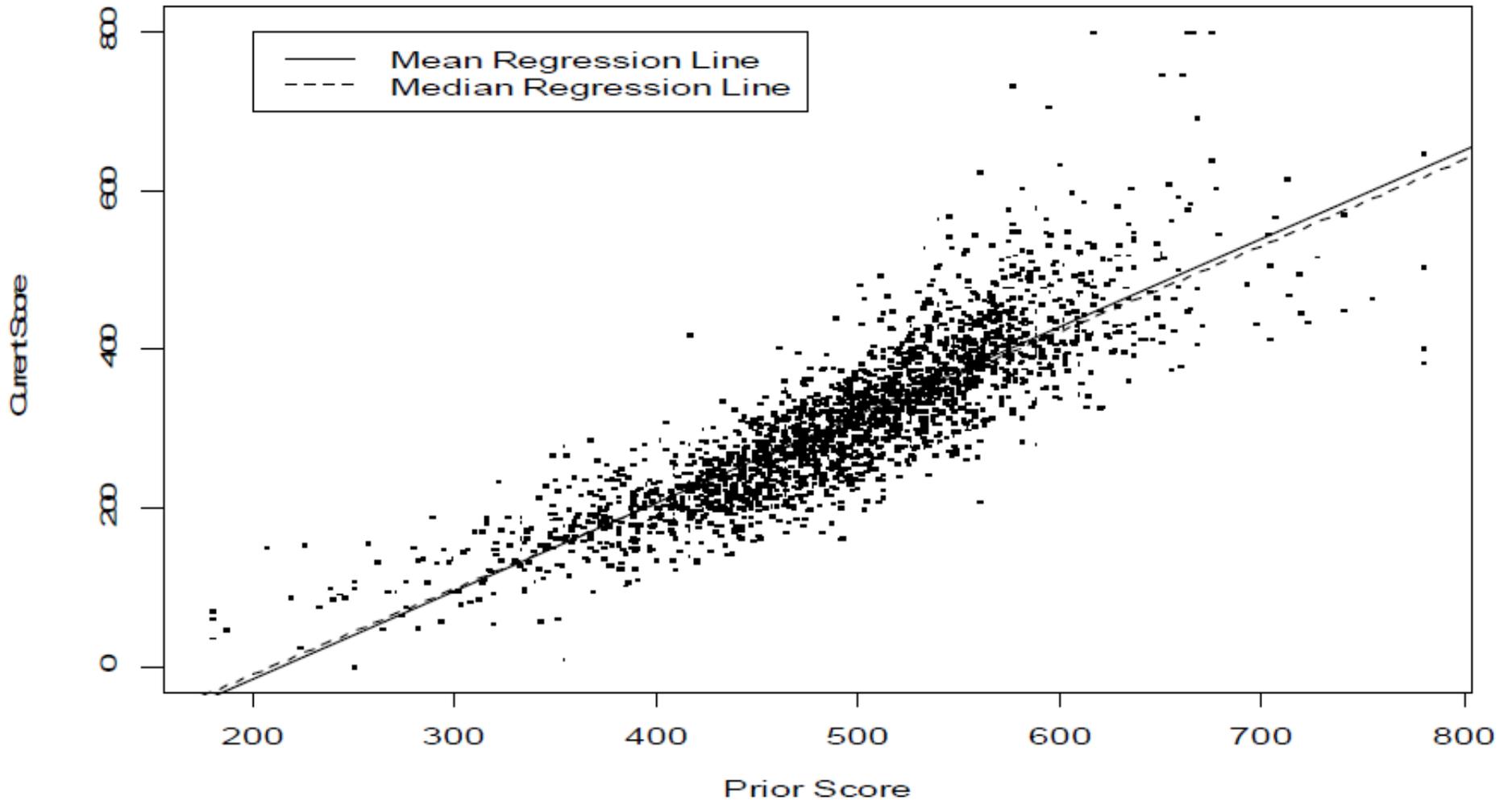
- Controls for the inability to accurately estimate abilities at the extreme ends of the distribution;
- Clumps the data points at the ends of the distribution (specifically, when there's a floor and/or ceiling effect)

What does the scatterplot look like, when scores are translated to percentiles?



# From *Is Growth in Student Achievement Scale Dependent?*

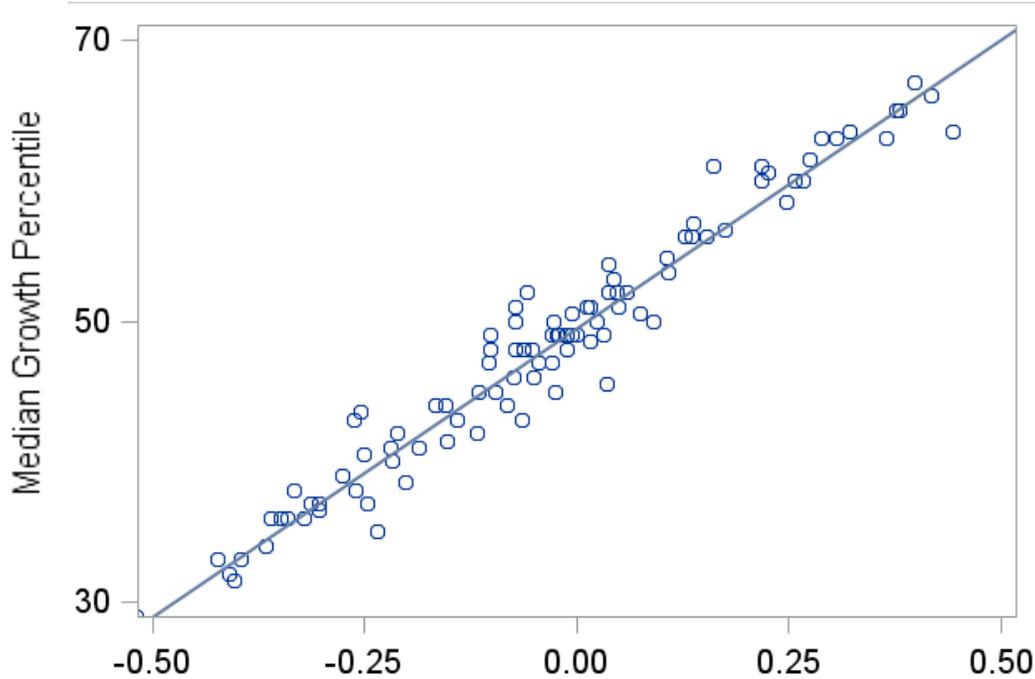
Derek Briggs & Damian Betebenner



# What's the relationship between Median Growth Percentiles and Linear Regression, at the **School** Level? (106 Schools)

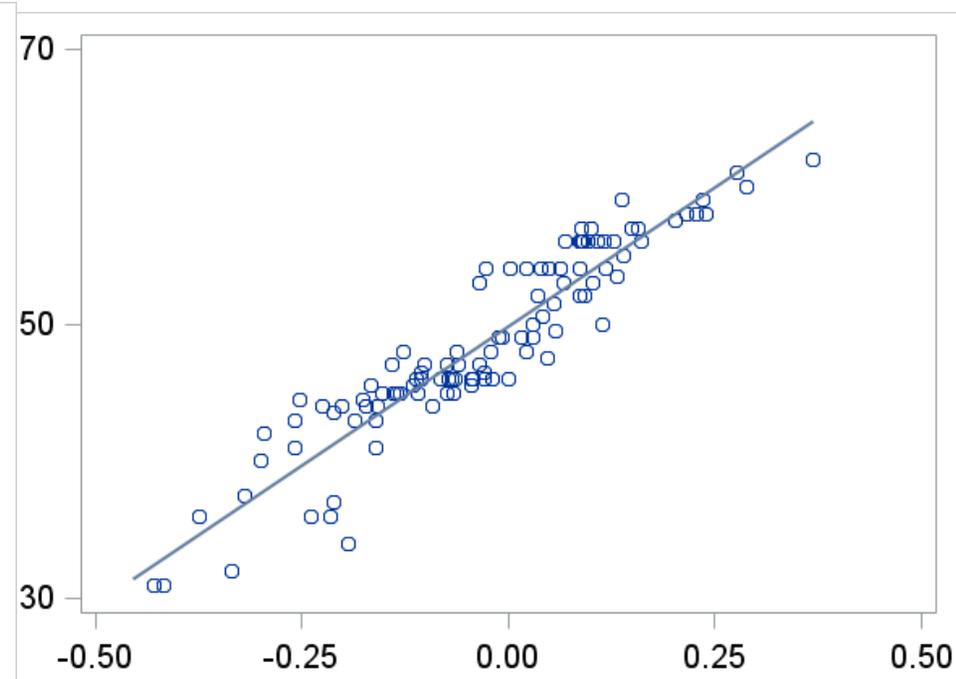
Math

$r = .98$



Reading

$r = .94$

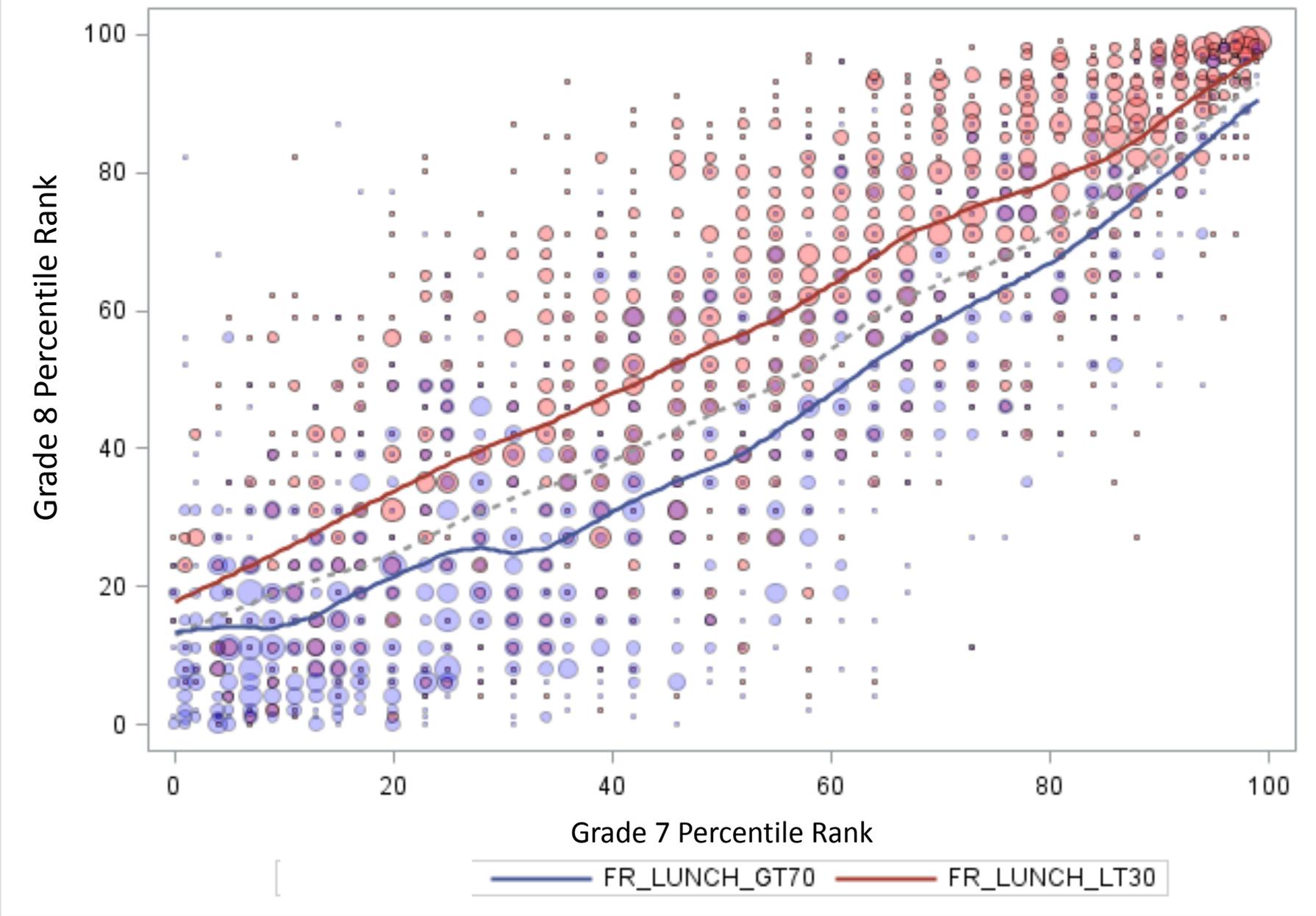


Linear Regression - Mean Standardized Residual (by-grade z-scores)

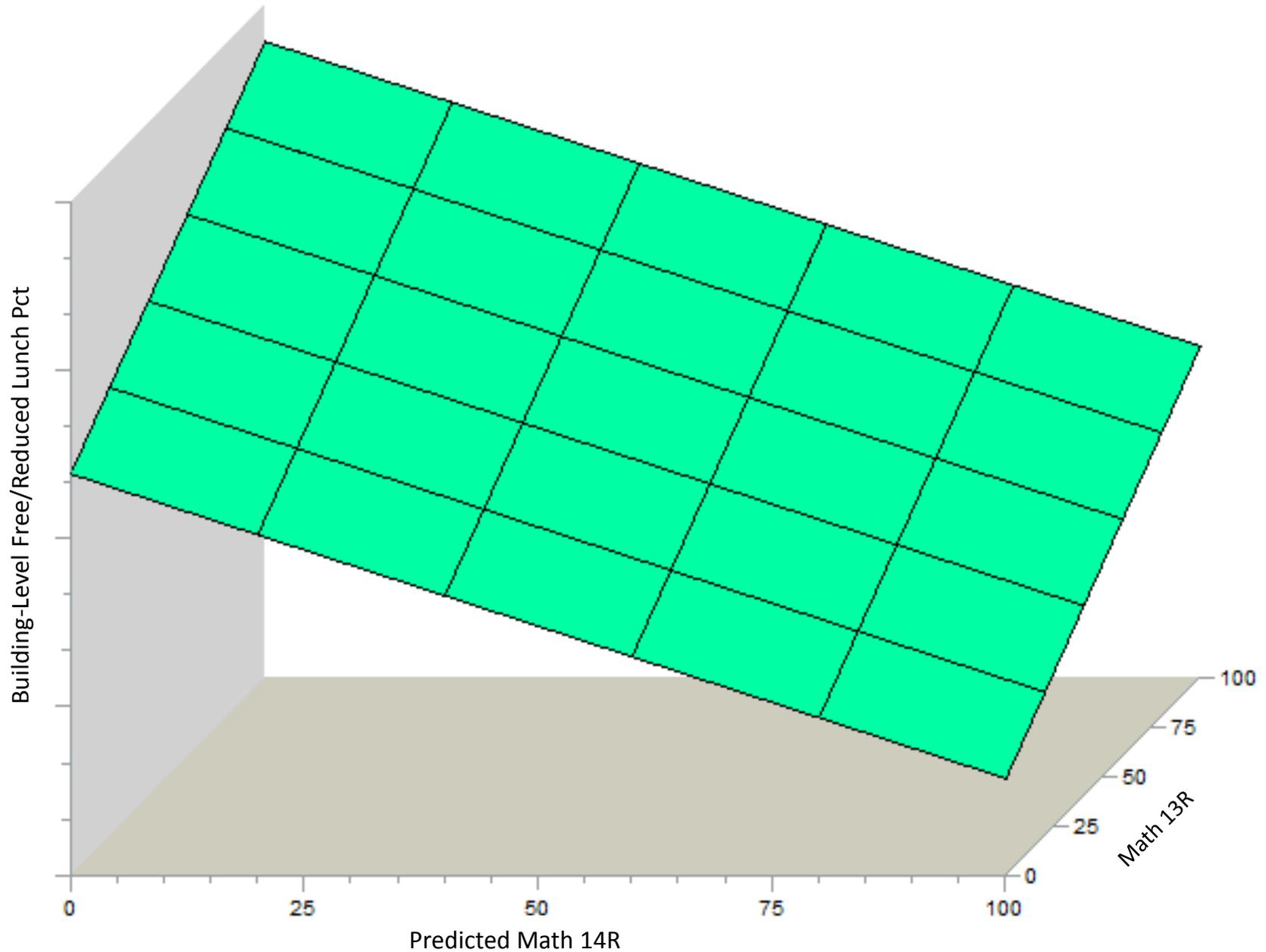
• Z-scores calculated from available Scale Scores, capped at -2/+2)

# Are the Median *Targets* Different, for High and Low Poverty Schools?

MEAP MATH, Grade 7 to Grade 8

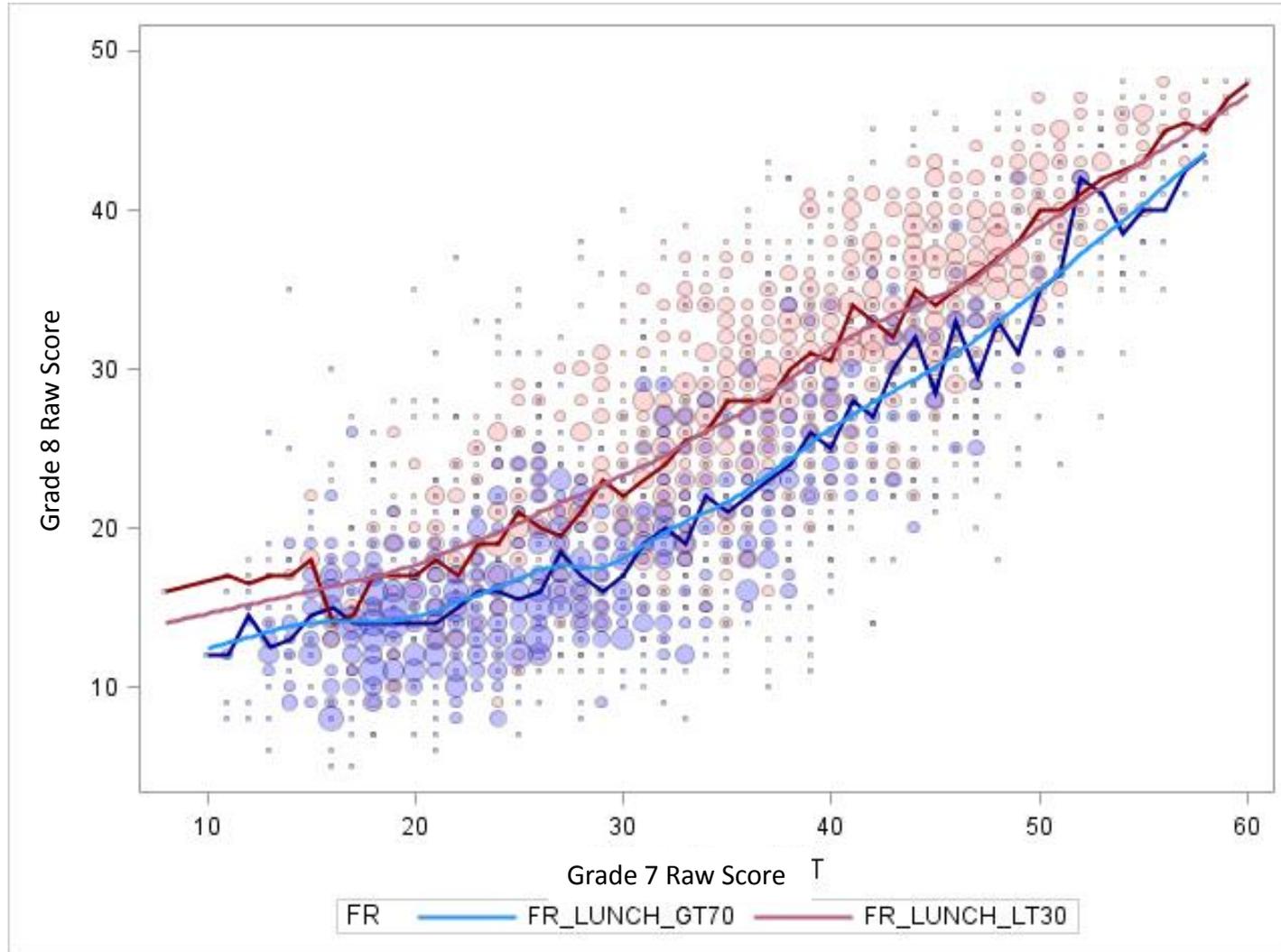


# A Scatter Plot, with 3 Dimensions



# Smoothed and Actual Median Percentiles – Switching to Raw Scores\*

MEAP MATH, Grade 7 to Grade 8



Lack of smoothness most likely do to sparseness of data (and we only have 60 point values).

\*Raw Score – Scale Scores conversions were the same across forms (except form 99, and took these out)

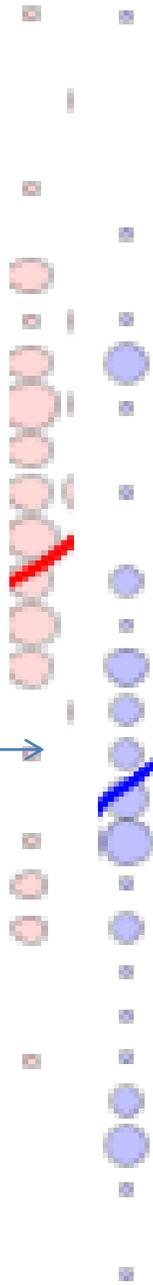
# What is the Difference Between the Two Groups of Students?

Median Pctile Rank

FR Lunch  
Less Than 30%  
- 30 pts



FR Lunch  
Greater Than 70%  
- 24 pts



48 possible Points

->  $30/48 = 63\%$  correct

->  $24/48 = 50\%$  correct

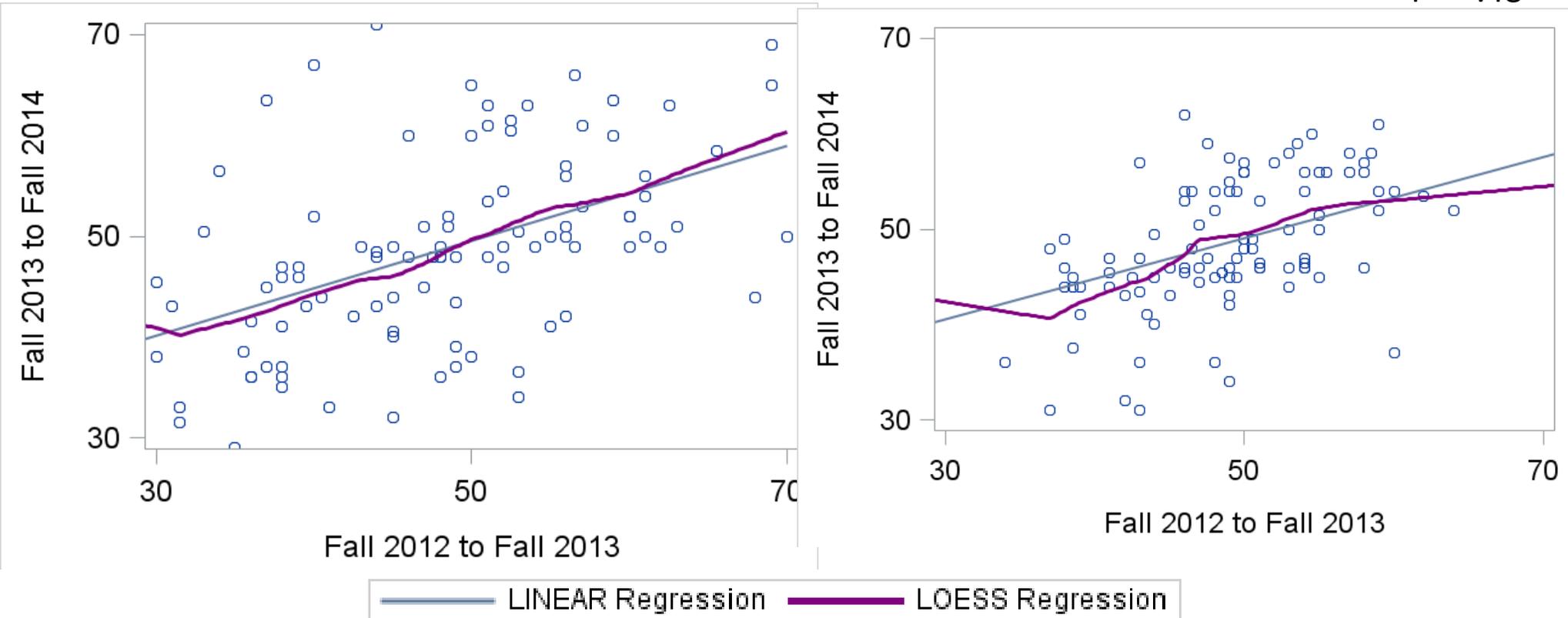
# How Consistent (stable/reliable) are School-Level Median Growth Percentiles?

## Math

$r = .46$

## Reading

$r = .48$

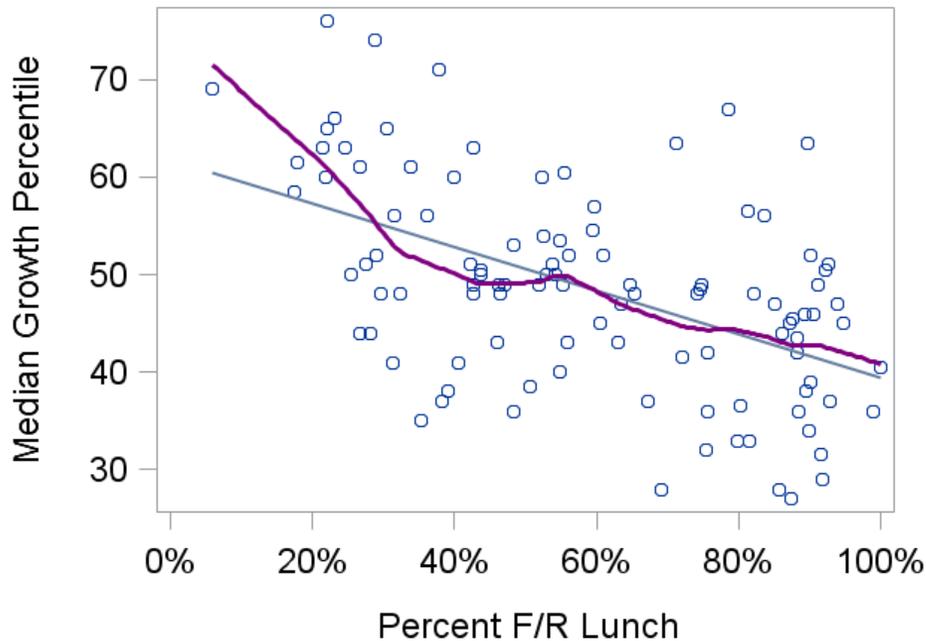


# And where does the Stability come from?

What's the relationship between School-Level Percent Free/Reduced Lunch and Median Growth Percentiles?

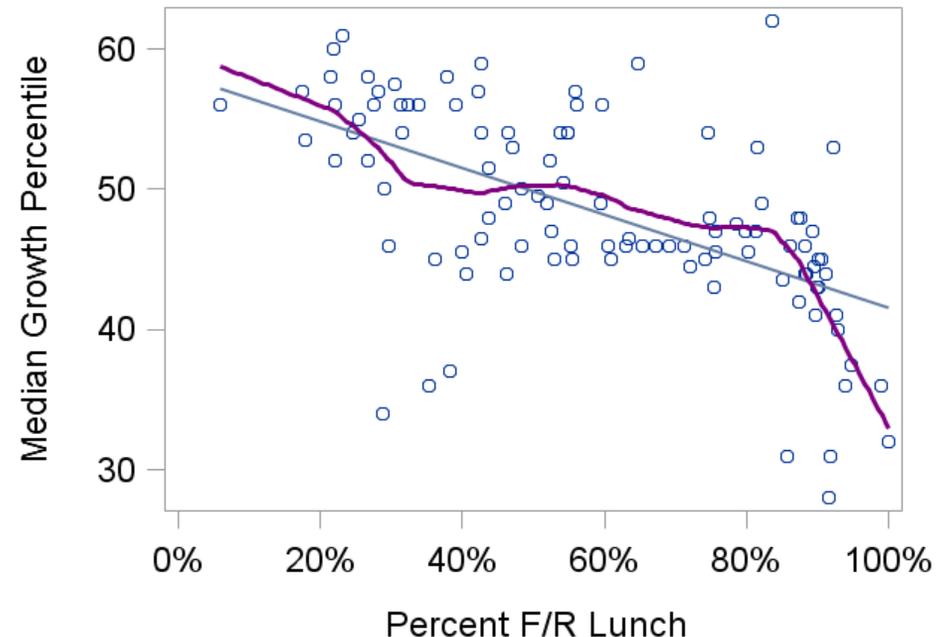
## Math

$r = - .52$



## Reading

$r = - .58$



— LINEAR Regression — LOESS Regression

### From Is Growth in Student Achievement Scale Dependent?

Correlation between median SGP and percent free/reduced lunch, across the three year/grade analyses: ranged from -0.42, -0.25, to -0.39

# New Jersey: Year 1 to Year 2 School-Level SGPs

Same schools, 1yr later,  
LA SGPs correlated at .61

	LA SGP Year 1	Math SGP Year 1	LA SGP Year 2	Math SGP Year 2	Math SGP 2yr Avg	LA SGP 2yr Avg
Math SGP Year 1	0.6224					
LA SGP Year 2	0.6137	0.4825				
Math SGP Year 2	0.4656	0.6116	0.5816			
% Proficient Math 4	0.5066	0.4706	0.5066	0.3268	0.4491	0.5659
% Proficient Math 7	0.5589	0.5438	0.5656	0.426	0.528	0.6106
% Proficient Language 4	0.5803	0.407	0.5228	0.3337	0.4159	0.6163
% Proficient Language 7	0.5739	0.4483	0.5888	0.3593	0.4392	0.6325
% Free Lunch	-0.4832	-0.3704	-0.5559	-0.3164	-0.3813	-0.5781

2yr avg. SGPs nearly as strongly (negatively) correlated with % free lunch as year over year SGPs with themselves!

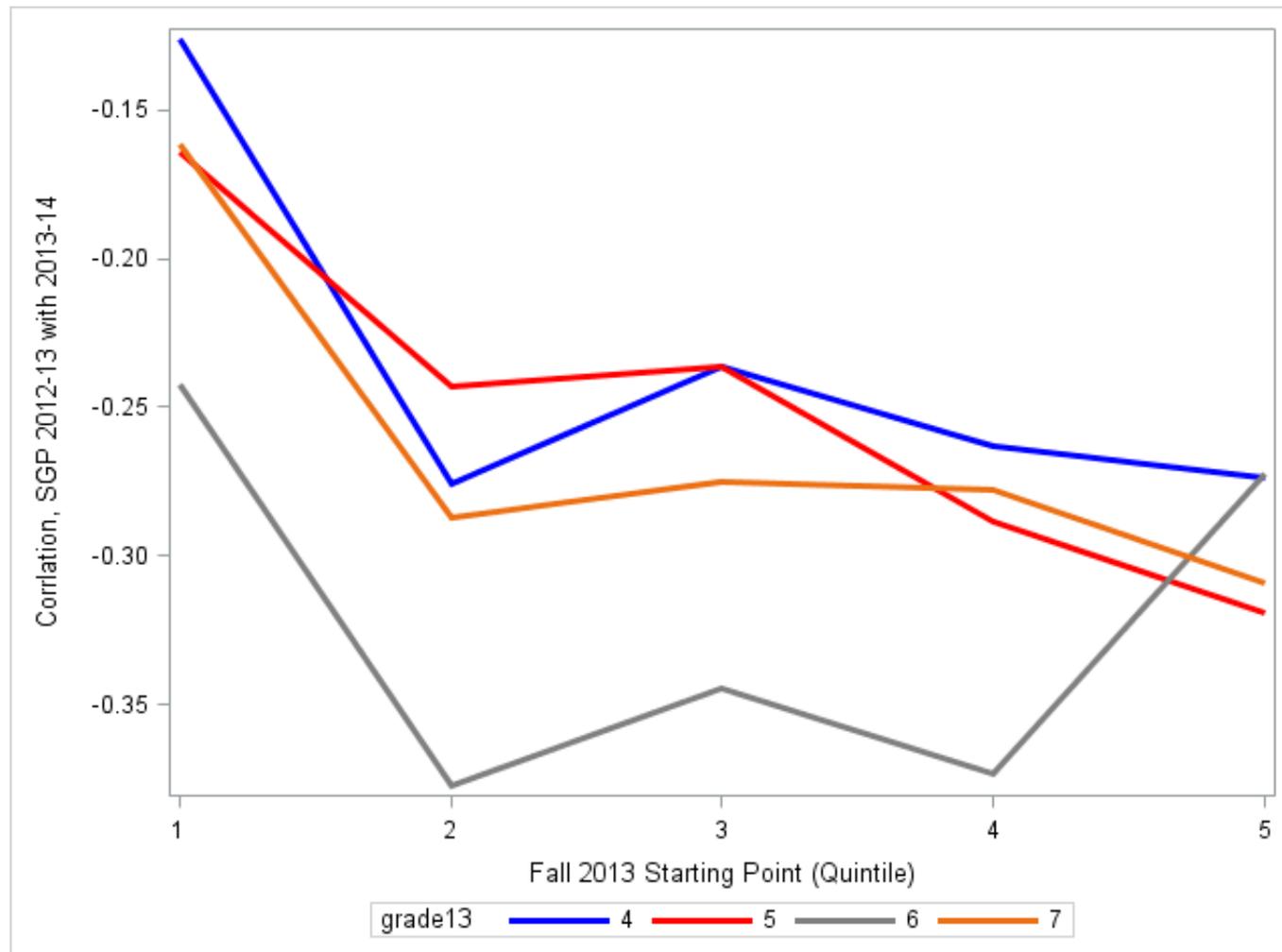
Avg. year 1 proficiency rates more correlated with 2yr avg. SGPs than year over year SGPs with themselves!

Sometimes I make graphs that don't work so well...



# Relationship between Previous and Current Year SPGs

Correlation, between 2012-13 SGP and 2013-14 SGP, By Grade and Starting Point



- Take away: in general, kids who had had a higher SGP, during "year 1 to 2" will have a smaller SGP in "year 2 to 3" (and vice versa)

# A Couple Points From the Article *Is Growth in Student Achievement Scale Dependent?*

by Derek Briggs & Damian Betebenner

- Student growth percentiles (SGPs) are computed for the students in our longitudinal cohort separately by grade ... [with] grade 6 conditioned on grades 3, 4 and 5.
- These SGPs are then aggregated to the school-level by taking the median. We do not refer to school-level SGPs as value-added estimates for two reasons. First, no residual has been computed (though this could be done easily enough by subtracting the 50th percentile), and second, ***we wish to avoid the causal inference that high or low SGPs can be explained by high or low school quality.***  
(I added emphasis)
- The correlations between median SGP and current achievement are (tautologically) higher reflecting the fact that students growing faster show higher rates of achievement that is reflected in higher average rates of achievement at the school level.

# Ever see a quote and say "Wish I said that"?

Let's just say I am skeptical that either VAMs (Value Added Models) or SGPs (Student Growth Percentiles) can provide useful insights to anyone who doesn't have a pretty good understanding of the nuances of these kinds of data/estimates & the underlying properties of the tests.

If I was a principal, would I rather have the information than not? Perhaps. But I'm someone who's primary collecting hobby is, well, collecting data. That doesn't mean it all has meaning, or more specifically, that it has sufficient meaning to influence my thinking or actions. Some does. Some doesn't. Keeping some of the data that doesn't have much meaning actually helps me to delineate.

Bruce D. Baker  
Rutgers University

Thanks!