Reaction to

Assumptions of Value-Added Models for Estimating School Effects

- Sean F. Reardon
- Steven W. Raudenbush

Joseph Martineau, Interim Director
Office of Educational Assessment & Accountability
Michigan Department of Education

Global Reactions

- In theory, theory and practice are the same. In practice, they are not
  - Lawrence Peter Berra
  - http://www.quoteworld.org/quotes/1305
- The reality at hand
  - Theory: assumptions are met
  - Practice: “assumption is not plausible, but it is hard to conceive of the value added project going forward without it” (pp 33-34).
- This paper responds well to a conflicted reality
  - A strong model for the evaluation of the fitness of VAM for practical use
  - Theory says assumptions must be met
  - Practice asks what degree of violation is acceptable

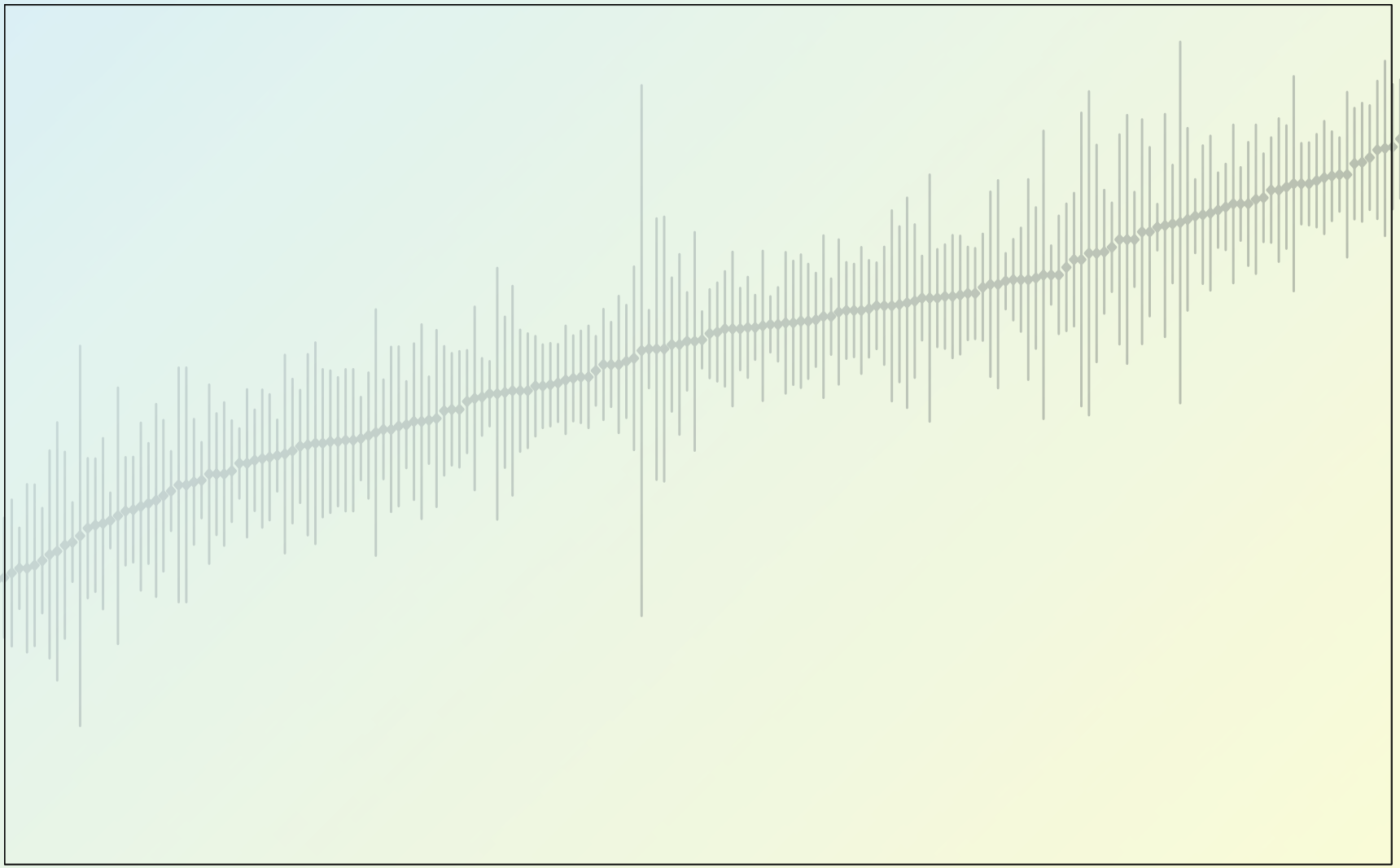
Specific Reactions: First Three Assumptions

- These assumptions are studied by other authors, but…
  - Accepting that assumptions do not hold
  - How do violations of first three assumptions affect correlations in presence of violations of the second three assumptions?
  - Presents a best case scenario
  - Unclear how important interactions are

Specific Reactions: Second Three Assumptions

- Good news!
  - Polynomial forms help
- Question
  - How was range of curvature selected?
- Important missing “factor”
  - Dimensionality

Curvature, Empirically



Scale Compression? Discrepancy in Ranks?

A “non-grotesque” transformation?

Curvature, Limitations

- Non-linear transformation of a unidimensional measure
- Psychometric claims of interval level measurement
  - Interval measurement arguments rest on unidimensionality
  - Unidimensional measures are based on the first principal component
  - A strong first principal component does not indicate unidimensionality (think height and weight)
- A first principal component of multiple manifest variables…
  - Is not a scale
  - May approximate a composite scale if curvature is modest
  - Under significant curvature, represents a non-linear trajectory through multidimensional achievement space

Curvature: Limitations, continued…

- Curvature (and scale compression) are important, but…
- Underlying dimensionality is more important
  - With a composite outcome
    - Assumes the same effect on all dimensions
    - Assumes the "unidimensional" scale linearly traverses multidimensional achievement space
- The interaction of non-linearity and multidimensionality is troubling
- Statistical fixes work for unidimensional data
  - Ordinal data, ranking of "more or less" learning impact
  - Polynomial functional forms
- Vertical articulation of expectations
  - Social moderation of "equivalent" points on adjacent grades?
  - How much does this address the issues of multidimensionality?

Brief Survey…

- Are weight and height manifest variables governed by a single underlying latent trait (e.g. “size”)?
- Are items measuring physical science, biology, and application of the scientific method manifest variables governed by a single underlying science understanding?
- Are computation and algebra items manifest variables governed by a single underlying mathematics understanding?
- Are decoding and comprehension manifest variables governed by a single underlying reading comprehension?
- How about comprehension of narrative versus informational text?

Brief Survey, continued…

- If composite slope is greater for 1st graders than 6th graders, do 1st graders…
  - gain more in size than 6th graders?
  - learn more math than 6th graders?
  - learn more reading than 6th graders?
  - learn more science than 6th graders?

Brief Survey, continued…

- Where are we comfortable claiming unidimensionality knowing that a strong first principal component is insufficient?
- Are we comfortable assuming that teachers have the same impact on…
  - *decoding* and *comprehension?*
  - *algebra* and *computation?*
  - *physics* and *application of the scientific method?*
- More robustness studies are needed in the vein of this study to address both statistical and measurement issues