

# Minimal Requirements for Robust Value-Added Estimates of Teaching Effectiveness

William L. Sanders  
SAS EVAAS  
SAS Institute, Inc.

# Test Data Requirements:

---

1. Scales of measure must be highly correlated with curricular objectives.
2. Scales must have sufficient stretch to allow the measurement of academic progress for both low and high achieving students.
3. Scales must have appropriate reliabilities.
4. Scales must have been constructed so that the expectation of progress with the same schooling effectiveness is the same regardless of the entering level of the student, OR the data must be transformed to meet this condition.

# Minimal Modeling Requirements:

---

1. Any models that use prior student test scores as predictor variables must use at least three prior scores. If not, the errors in the predictor variables will result in serious biases in the estimation of teaching effectiveness.
2. Even when three or more prior scores are available, many students will have varying amounts of prior scores. The modeling must be sophisticated enough to allow all the data for each student to contribute to the estimates.
3. The modeling must accommodate various teaching assignments, such as self-contained classrooms, team teaching and departmentalized instruction.
4. The teaching effectiveness estimates must be shrinkage estimates (BLUP) to mitigate both false negatives and positives merely due to the luck of an accumulation of random errors.

# Characteristics of Our Preferred Teacher Effectiveness Model:

---

Multivariate (all subjects simultaneously), longitudinal analyses

- Dampens the selection bias problem to triviality
- Dampens the problem of errors of measurement in prior scores to triviality
- Past, present and future data are included simultaneously
  - Provides protection against testing irregularities
  - Provides a dampening of unexpected pulses that can affect performance during testing, i.e. tornado alert, etc.

# Models That *Never* Be Used for Teaching Effectiveness Measures:

---

- Simple paired mean gain averages
- Mean residuals from simple regression models
- More complicated models that do not include the data for each student with missing test scores (results in a non-trivial selection bias problem)
- Any model without shrinkage estimates of the classroom effects

## In Sum:

---

- To trade reliability of information for simplicity of calculation is a “Devil’s Bargain.”
- Results from simplistic value-added teaching models will produce an alarming number of false positives and negatives.