Technical and Practical Considerations in applying Value Added Models to estimate teacher effects

Pete Goldschmidt, Ph.D.

Educational Psychology and Counseling, Development Learning and Instruction

California State University Northridge
Michael D. Eisner College of Education

Purpose VAM

- What motivates the use of Value Added Models?
  - Improve student performance
  - Teacher pay/licensure decisions
- School Accountability
  - Free rider problem
- Teacher Evaluation
  - Elements of teacher evaluation
    - Inputs as indicator
    - Observations as indicators

VAM Introduction

- Vocabulary
- What are VAMs trying to measure?
  - Teacher contribution to student learning
    and attempting to make this a causal estimate

- Is there a plausible alternative? Unlike experimental setting where plausible alternative is specified, use average as basis for comparison.
  - Average of what?
  - School
  - District
  - State

Technical considerations

- Bias
  - Student sorting
- Precision
- Reliability
- Stability

California State University Northridge
MICHAEL D. EISNER COLLEGE OF EDUCATION

Practical considerations

- Assessments and scale
- Available data and linkages
- Tested and non-tested subjects
- Components of evaluation system

Models

- Models vary in complexity and assumptions
- Models vary in application
  - Education basis
    - generally random effects models
  - Economist basis
    - Generally fixed effects models
  - State accountability based models

A First Approximation

- $A_{ti} = X\mathrm{B} + S\Gamma + T\Phi + e_{ti}$

- X = vector of student and family inputs
- S = vector of schooling inputs
- T = vector of teacher inputs
- Assumes e_{tj} is orthogonal to covariates, which is highly unlikely

California State University Northridge MICHAEL D. EISNER COLLEGE OF EDUCATION

Rationale Underlying Value Added Models

- The underlying assumption for value added models is:

$$A_{it} = f(B_{it}, P_{it}, S_{it}, I_{it}, E_{it}), \qquad (1)$$

- where for student *i* at time *t* Achievement A, is some function of:
  - *Student background (B)*
  - *Peer and other influences (P)*
  - *School/teacher inputs (T/S)*
  - *Innate/general ability (I)*
  - *And luck (E).*
- Model is cumulative and past inputs may affect current Achievement.
- Also would need independent measure of innate ability, gathered before any S has occurred.

Specification

- If we assume that (1) holds for any time t, then we can consider change in achievement from t to t`.

  $A_{it`} - A_{it} = f(.)$

VAM Specified

- Simplified specification

$$A_{it} = \delta A_{it-1} + T_{tj} + \alpha_i + e^*_{it}$$

where $e^*_{it} = e_{it} - e_{it-1}$

T_j is teacher j’s effect and

α_i is an individual student time-invariant effect.

If assume $\delta = 1$ then use common “gain” model

$$A_{it} - A_{it-1} = T_{tj} + \alpha_i + e^*_{it}$$

California State University Northridge MICHAEL D. EISNER COLLEGE OF EDUCATION

Assumes

- age independence
- additive separability
- fixed family inputs
- geometric decay in previous inputs
- homogenous teacher effectiveness
- sorting based on fixed student covariates
- OLS produces biased estimates because A_{t-1} and that part of e* related to e_{t-1} - not orthogonal.

California State University Northridge
MICHAEL D. EISNER COLLEGE OF EDUCATION

Student Effects

- Student fixed effect α_i can be modeled by:
  - Dummy variables
    - Empirical evidence suggests that this biases teacher effects downwards.
  - Using student time invariant covariates
  - Assuming adequately captured by A_{t-1}
  - Use instrumental variables or additional test scores for A_{t-1}

California State University Northridge MICHAEL D. EISNER COLLEGE OF EDUCATION

Teacher effects

- Can estimate T_j by
  - Using dummy variables for each teacher
  - Demeaning by teacher means
    - Generating within unit estimates

Fixed effects

- Avoids bias (if sorting based on static factors)
  - Some evidence that sorting is based on dynamic factors.
- Limitations
  - Can't estimate time invariant effects
  - ignores between teacher variability
    - Teacher effects will be less precise than when using random effects models

Random Effects Models

- Common approach used by educational researchers/statisticians
- Can be residualized gain, or growth specification.

ANCOVA Specification

- $A_{tij} = \delta A_{t-1ij} + \phi T_j + e_{tij}$

- Where ϕ is a random teacher effect where:

  $\phi = \gamma + U_j$

  Hence $A_{tij} = \delta A_{t-1ij} + \phi(\gamma + U_j) + e_{tij}$

  and U_j ~N(0,τ) and assumed orthogonal to student covariates that may be in the model

  Generally use EB estimates which “shrink” estimates towards mean (effect depends on reliability of teacher effect estimate).

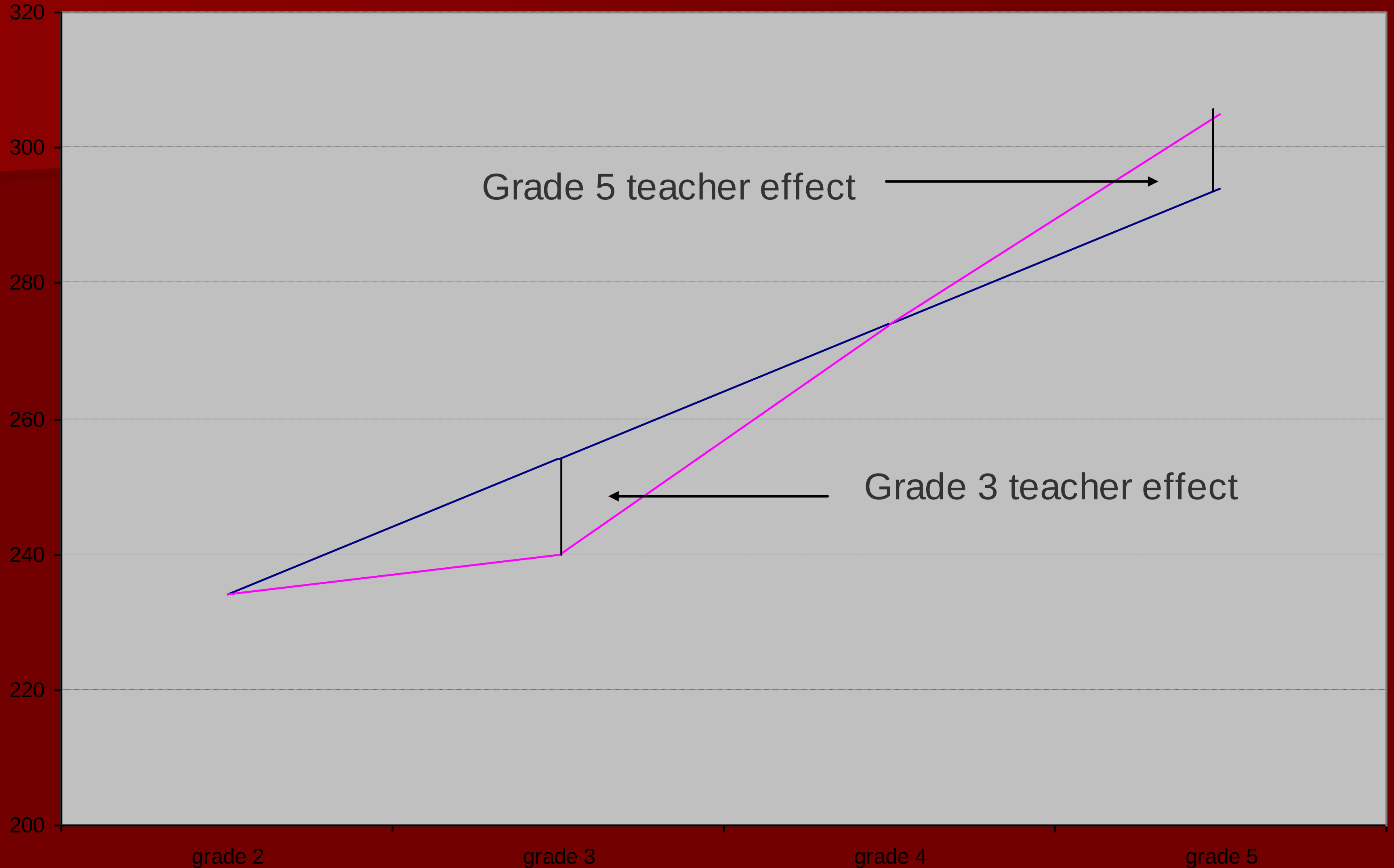
Random Effects tradeoff

- Random effect teacher estimates are more precise
- Random effects may be biased if more restrictive (compared to fixed effects models) assumptions are not met.
- Random effect models can include time invariant covariates.
- Models can be "centered" around group means to recreate fixed effects estimates.

Growth Model Approach

- Can model both A_t and A_{t-1} on the LHS of the equation avoiding correlations of error and A_{t-1} or other covariates.
- Can model multiple assessment occasions over time as a function of time.

Conceptualization of teacher effect as deviation from average trajectory



California State University Northridge
MICHAEL D. EISNER COLLEGE OF EDUCATION

Estimating True Status and Gain

(an example of an optional approach, based on*)

- $A_{tij} = a_{1tij}\pi_{1ij} + a_{2tij}\pi_{2ij} + e_{tij}$

- Here, for student *i*, with teacher *j*, at time *t*, the assessment scale score is denoted as A_{tij}. Time in this instance refers to the A_{t-1}, $t = 0$, and the A_t $t = 1$.
- This estimates two parameters, student's initial status for the A_{t-1} (π_{1ij}) and gain on the A_t, (π_{2ij})

*Bryk, A., Thum, Y. M., Easton, J., & Luppescu, S. (1998). Assessing school academic productivity: The case of Chicago school reform. *Social Psychology of Education*, 2, 103p142

California State University Northridge — Michael D. Eisner College of Education

The error, et_{ij}, is assumed to be

$N \sim (0, \sigma^2)$

Can conceive as the student growth part of the model as a measurement model and incorporate precision (SEM) to identify the model.

- $A^{*}_{tij} = a_{1}{}^{*}_{tij}\pi_{1ij} + a_{2}{}^{*}_{tij}\pi_{2ij} + e^{*}_{tij}$
- In this way e^{*}_{ijt} is distributed $N \sim (0,1)$, and π_{1ij}, and π_{2ij}, now estimate a student's true initial status and true gain, respectively. Can estimate teacher effect using estimate of true gain.
- * indicates parameter is weighted by precision ($1/SEM_{ij}$).

- In order to replicate fixed effects estimates group mean center

- $A^{*}_{tij} = a1^{*}_{tij}\pi 1_{ij} + (a_2{}^{*}_{tij} - a_2{}^{*}_{.ij})\pi_{2ij} + e^{*}_{tij}$

- where $e_{tij} \sim N(0,1)$, and $a_2{}^{*}_{.ij} = S_t{=}1,2\ a_2{}^{*}_{tij}/2$ (2 in this case because we have a A_t and A_{t-1}).

California State University Northridge
MICHAEL D. EISNER COLLEGE OF EDUCATION

- *Hence, combined model is:*
- $A^{*}_{tij} = \gamma_{1tij}\pi_{1ij} + \gamma\pi_{2ij} + r_{ij} + u_j$

- where $r_{ij} = r_1 a_1^{*} + r_2(a_2{}^{*}{}_{tij} - a_2{}^{*}{}_{.ij})$,
  and $u_j = u_1 a_1^{*} + u_2(a_2{}^{*}{}_{tij} - a_2{}^{*}{}_{.ij})$.
- Uj is estimated teacher effect

California State University Northridge MICHAEL D. EISNER COLLEGE OF EDUCATION

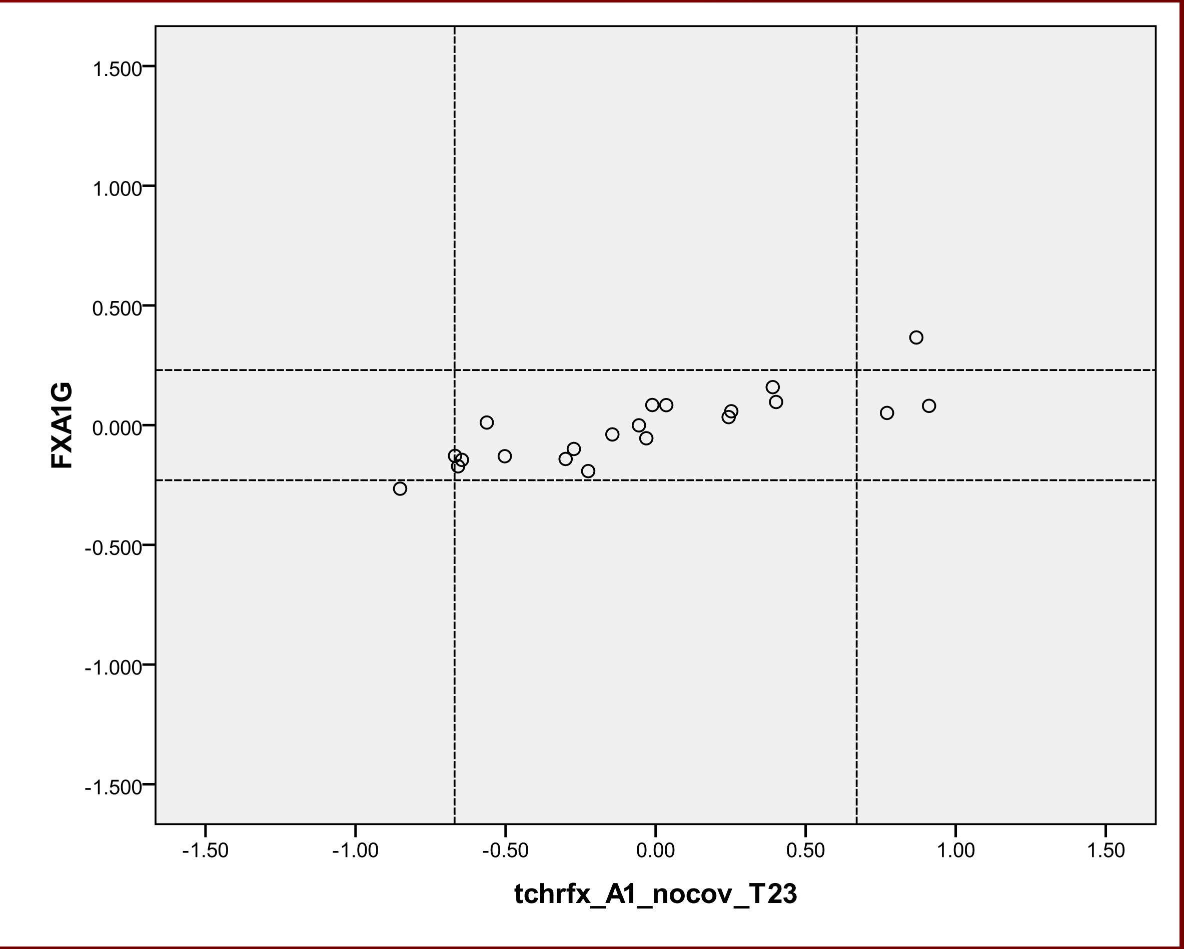
- Advantage of including student covariates:
  - Can estimate initial achievement gaps
  - Can estimate time to close gap
  - These estimates likely too imprecise at the teacher level, but are useful at the school level

Technical considerations

- Bias
  - Sorting
  - Measurement error
  - Equating error
- Precision
  - Classification categories– e.g. highly effective
  - Better precision than existing teacher evaluation measures?
    multiple assessments
- Reliability
  - Ability to detect true between-teacher difference
- Stability
  - Classifications stable over time
  - Multiple year (congruent with policy?)

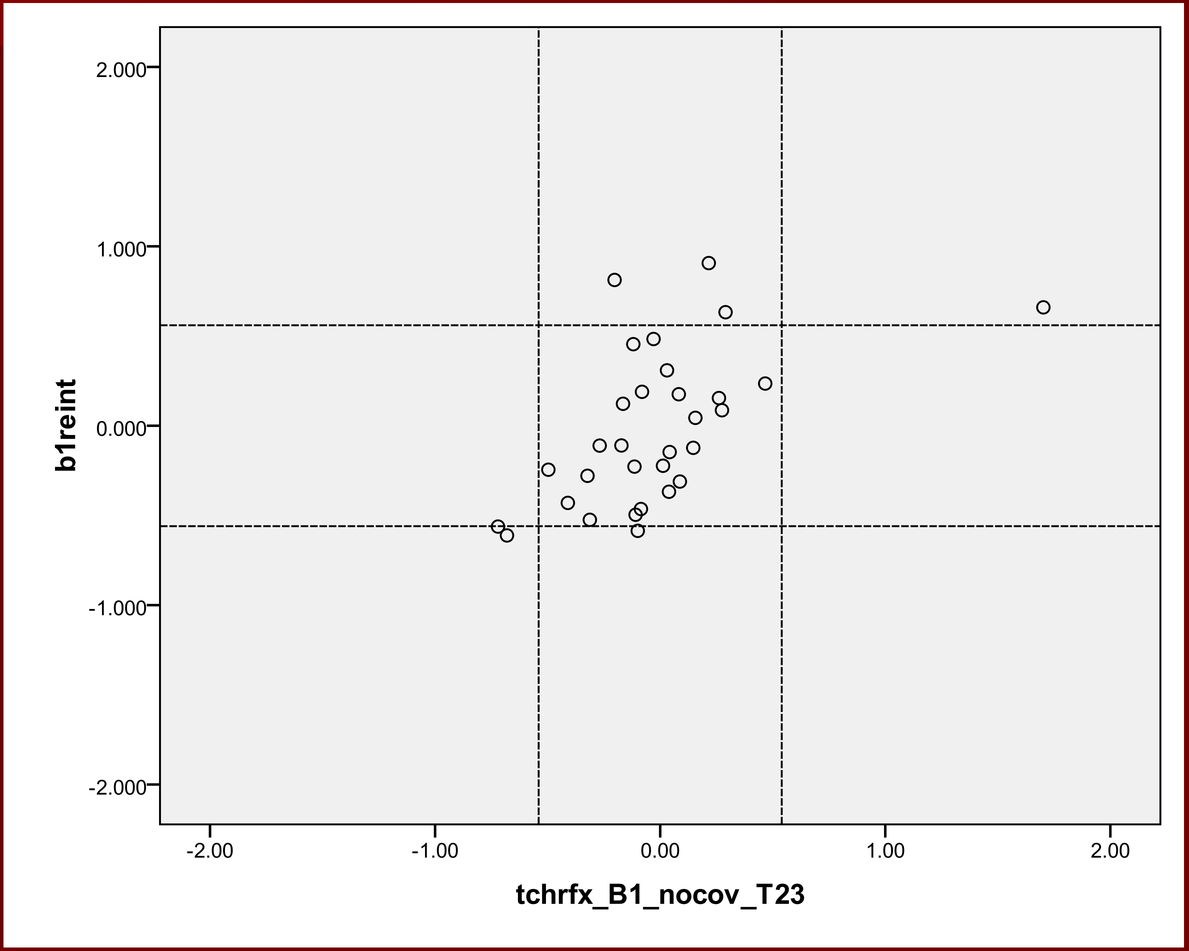
California State University Northridge — Michael D. Eisner College of Education

Effect of Model and Precision



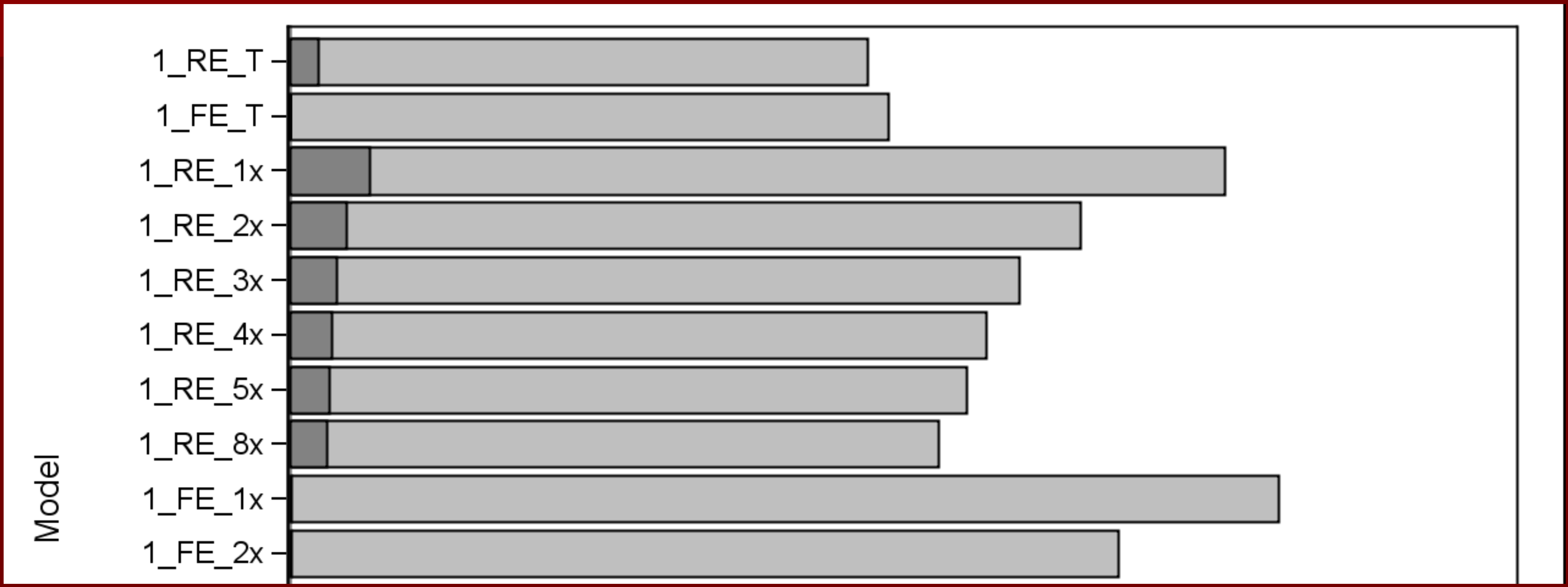
California State University Northridge
MICHAEL D. EISNER COLLEGE OF EDUCATION

Effect of Model and Precision



California State University Northridge
MICHAEL D. EISNER COLLEGE OF EDUCATION

Effect of measurement error



From Wright, P, 2008

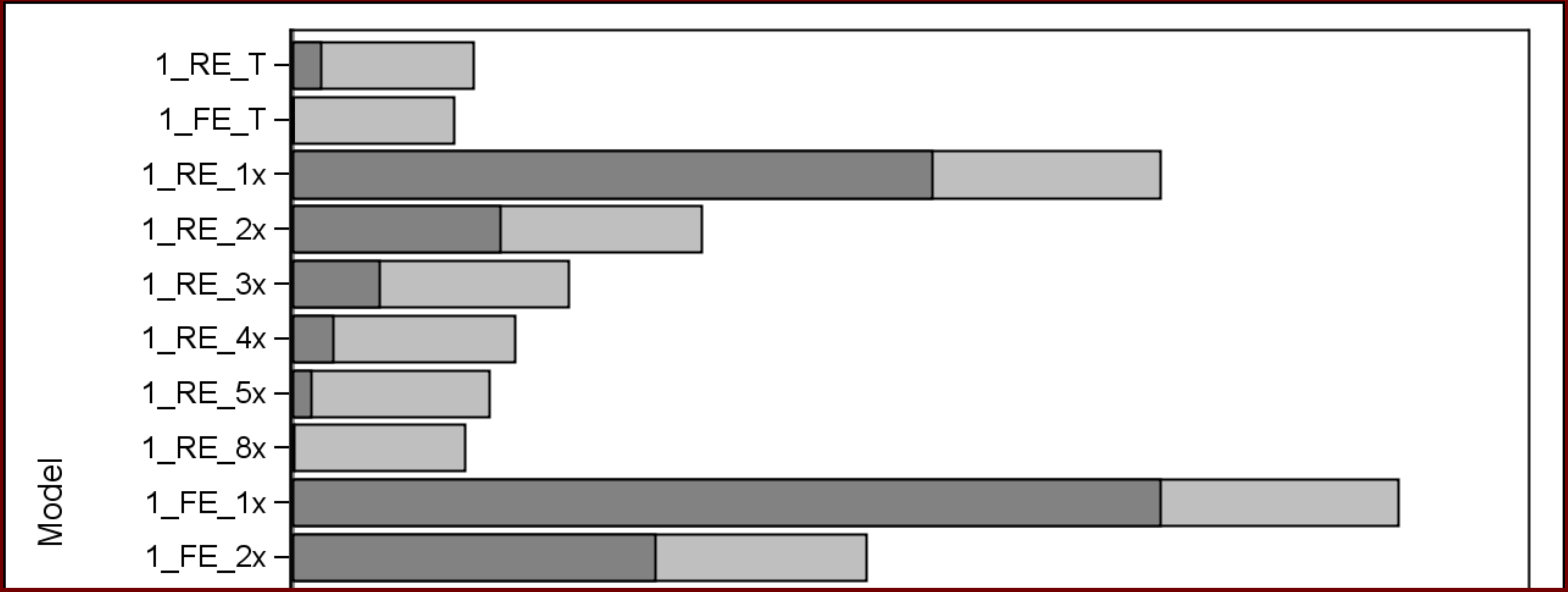
Results with no measurement error in assessment

Light gray = se

Dark gray = bias

Measurement Error and Sorting bias

- Overall, half of the students were free/reduced-price lunch eligible.
- For individual students, Prob(Pij=1) decreased with increasing true pre-test score (ξi).
- Lower performing students (lower ξi) were more likely to be assigned to a poorer teacher.



Results with measurement error in assessment

From Wright, P, 2008

Light gray = se

Dark gray = bias

California State University Northridge MICHAEL D. EISNER COLLEGE OF EDUCATION

Practical considerations

- Assessments and scale
- Available data and linkages
  - Spill over
  - Persistence
- Tested and non-tested subjects
- Components of evaluation system

Pete.Goldschmidt@csun.edu

818-677-4601

California State University
Northridge
MICHAEL D. EISNER
COLLEGE OF EDUCATION