

The National Conference on Value-Added

UW-Madison, April 22-24

Program Chairs

Douglas Harris, Adam Gamoran, Steve Raudenbush

Program Committee Members

Tim Sass, Rob Meyer, Henry Braun, JR Lockwood

Funders

Carnegie Corp, Joyce and Spencer Foundations

Why Are We Here? (Part I)

- Student achievement is an important objective of schools and a tool in education policy
 - Beginning in early 1990s, it became common to use snapshots of student achievement to measure and reward school improvement
 - NCLB has strongly reinforced shift toward “accountability”
- Because student achievement is an important objective, researchers use this as an important outcome variable in program evaluations
 - The basis for education production function studies dating back at least to the Coleman Report

Why are We Here? (Part II)

- Common uses of student test scores for accountability and program evaluation are hindered by **selection bias**:
 - (1) Students are non-randomly assigned to teachers
 - (2) Schools and classrooms are non-randomly assigned to programs
- In theory, value-added addresses these selection bias problems

An Important Distinction

- As we will see, the “right” value-added model depends on how it is being used—that is, on the types of conclusions being drawn
- Value-added for program evaluation (VAM-P)
- Value-added for accountability (VAM-A)

Encouraging Observations about VAM-A

- Differences between the lowest and highest value-added teachers seem large
- Some evidence that having a high-value-added teacher has lasting benefits for students
- VAM-A measures of teacher effectiveness are positively correlated with principals' subjective assessments of teachers
- VAM-A measures have been validated in a random assignment experiment

Less Encouraging Observations

- Measured “teacher effects” are imprecise
 - hard to say that one teacher is clearly better than another based on VAM-A
- Related point: teacher effects are unstable
- Value-added of individual teachers varies--sometimes considerably--depending on how the model is specified

Possibly Problematic Assumptions

- VAM requires many strong assumptions, such as:
 - test scale
 - assignment based on fixed student and teacher characteristics
 - each teacher equally effective with all types of students
 - constant rate of decay in effects of past inputs
- Violation of these assumptions may explain some of the less encouraging results

Critical Question

How do we reconcile the encouraging findings with the less encouraging findings and problematic assumptions?

Broader Conference Goals

- To create greater clarity about the properties of VAM-A and VAM-P, from a statistical standpoint
- To narrow the range of appropriate value-added models (may vary by VAM-A and VAM-P)
- To provide a sense of the policy implications of these technical findings for use of VAM-A

Critical Point

*We must view VAM-A and VAM-P
in comparison to the alternatives*

Conference Ground Rules

- Each presenter has 20 minutes, followed by 10 minutes for discussants, followed by Q&A
 - please hold questions until the end of each session
- Remember, we have researchers from a wide variety of disciplines and fields—try to avoid jargon and explain the intuition
- Remember our two purposes: technical issues and policy issues, VAM-A and VAM-P