

# The “Policy Validity” of Value-Added and Other Teacher Quality Measures

Douglas N. Harris  
University of Wisconsin at Madison

# Outline

- Recall, importance of comparing VAM-A with the alternatives
- Here, I will consider specific alternatives:
  - (1) teacher credentials
  - (2) others
    - students' status attainment
    - peer and principal assessments of teachers
    - providing raw student data to teachers
    - school-level value-added
- To make these comparisons, we need a framework

# “Policy Validity”

- Policy validity refers to the appropriateness of the use of the measures, based on:
  - (1) statistical validity
  - (2) function
  - (3) cost
- Some measures with low statistical validity still have “valid” policy uses
- Some measures with high statistical validity can be used in “invalid” ways

# Functions of Quality Measures

- Signaling/filtering involves determining who starts off in the profession
  - examples: certification; entry requirements to teacher education programs
- Improvement involves raising effectiveness for those who have started in the profession
  - type #1: summative, to recognize/reward success and failure
  - type #2: formative, to identify a path to improvement

# Functions of VAM-A

- Primary function is summative assessment
  - determine how well teachers are performing, their causal impacts—can be used for teacher tenure, compensation, etc.
- Could also play a signaling function
  - anecdotally, some school principals in Tennessee require prospective teachers to bring their VAM information to interviews
  - states could create a clearinghouse that school principals (and districts) could access

# Evidence on Teacher VAM-A

- We are making good progress
- Overall, still seems that value-added is informative, but that many of the assumptions and statistical properties are problematic
- See forthcoming conference summary paper
- See future work by this group

# Functions of Credentials

- Most credentials could serve both roles; however, with many credentials, nearly impossible to estimate the improvement (causal) impacts in many cases
- The signals are those credentials that are created before teachers enter the classroom:
  - undergraduate education, certification, tests (e.g. Praxis)
- Some are primarily for improvement:
  - experience, graduate education, prof.dev.

# Evidence on Credentialing

<i>Teacher Characteristic</i>	<i>Gain Score Studies</i>		<i>VAM-P/Experiment</i>	
	<i>Pos/Sign</i>	<i>Insignif., Neg.</i>	<i>Pos/Sign</i>	<i>Insignif., Neg.</i>
Undergraduate	5	4	0	3
Graduate	3	10	3	6
Prof. Develop.	0	1	2	1
Experience	7	8	8	1
Test score	5	2	1	1

# Policy Validity Depends on the Function

- It might make sense to give preference to teachers with master's degrees when hiring (signaling function), but not to include the master's degree in the salary schedule (improvement function)
  - especially for teacher-leader positions
- Conversely, it might not make sense to hire teachers who have taken more professional development, but still to reward teachers for taking PD after they're hired

# Costs of Teacher Quality Measures

- Often neglected topic in all areas of education
- Some measures much more costly than others:
  - teacher test scores and VAM-A are cheap
  - a master's degree for a teacher costs at least \$80,000 in budgetary and non-budgetary costs
- Need to consider the costs and benefits no matter which function a measure is used for

# The Policy Validity of VAM-A versus Credentials

- VAM-A has greater statistical validity as a signal and summative assessment
  - almost has to be true; credentials aren't intended to measure contributions to student test scores
  - partial exceptions: experience, some forms of professional development (possibly others, see NC & NYC studies)
- But, credentials provide a path to improvement
  - VAM-A does not
- VAM-A is cheaper than nearly all credentials

# Other Alternatives

- VAM-A almost certainly better than “status attainment”
- School VAM-A doesn't provide information about how each teacher is doing, but:
  - (a) school VAM-A creates pressures; and
  - (b) perhaps “everybody knows” who the good teachers are, creating peer/administrator
- Provide raw student sub-scores to teachers (along with school/district averages)
  - provides summative measures and path to improvement
  - but no selection bias adjustments; or incentives

# Conclusions

- In deciding how to use VAM-A, we must:
  - (1) compare it to the alternatives; and
  - (2) consider statistical validity in relation to specific functions; and costs
- Despite the problems, teacher VAM-A appears more cost-effective than credentials as signals and summative assessments
- However, we need a path to improvement, and existing credentials may serve that function
  - we know little about this, except experience

# Parting Thoughts (Part #1)

- Tempting to think “we should use whatever information we have”
- However, consider an organization with two objectives: achievement and “motivation to learn”
- Attaching high-stakes to only one measure could seriously distort behavior and reduce overall school success (a weighted sum of the two goals)
- For this reason, even if VAM-A had reasonably good properties (a “big if”), making high-stakes decisions without other measures is unwise
- Possible response: schools are so focused on things other than achievement that this is just correcting an existing distortion (rebuttal: NCLB)

# Parting Thoughts on Future Research

- Test value-added in use—in schools and school districts
- More randomized (and “apparently randomized”) experiments testing the validity of the models
- Continue testing validity of assumptions, especially combinations of assumptions
- But, focus more on robustness of models to modeling assumptions (e.g., correlation of teacher effects)
- Identify models that address the assumptions to which models do not seem robust