Selected Statistical Models for Incorporating Student Academic Growth in Estimating Effect of Teachers in Tested Grades and Subjects: Considerations Beyond Technical Adequacy

The expert panel appointed by the North Carolina Department of Public Instruction (DPI) has rated these three value-added models as having the technical adequacy needed for the high stakes purpose of estimating teacher effectiveness. WestEd was asked to help DPI consider more broadly the relative merits of each model, using criteria *other* than technical adequacy. This summary of "pragmatic" considerations is intended to support DPI when weighing tradeoffs associated with each of these models, assessing overall feasibility, and selecting the most appropriate model for statewide implementation.

Model	Face Validity with Stakeholders	Level of Expertise Required to Develop and Use	Ease of Implementation Statewide	Cost of Use/Resource Requirements	Policy Implications
3-Level Hierarchical Linear Model (HLM III)	May be perceived as more fair due to use of selected covariates (e.g., school SES) Used more frequently to estimate school effect	Model specification is complex and requires expertise With training and software license, analyses could be conducted by DPI Interpretation of findings requires expertise DPI would need to develop reporting strategies	Only model inclusive of grade 4 teachers Statewide cross-teacher comparisons may be problematic as school effect is not assumed to be constant across all schools; examining varying effects *within* schools can be meaningful Charlotte Mecklenburg Schools uses an HLM II but transition to HLM III likely to be smooth	Requires purchasing of HLM software package (annual license) but cost is not significant	When analyses are conducted appropriately, findings may be used as one component of NCEES and used for goal-setting at the student-, classroom-, school-, district-, and state-levels Impact will need to be monitored Stakeholder groups will need to be involved in determining cut-points for each effectiveness level
Univariate Response Model (EVAAS-URM)	May be perceived as stronger because NC adopted this model during NCLB GM Pilot and it has been widely used and adapted in states/districts across the nation	Model specification is complex and requires expertise With training and software license, analyses could be conducted by DPI Interpretation of findings requires expertise Reporting strategies and lessons learned can be borrowed from a	Excludes grade 4 teachers (2 yrs of prior data needed) Allows for statewide (cross-teacher) comparisons of teacher effect (school-level variables not required) Likely to be less challenging to implement statewide as a number of districts already are accessing results available	License is not required and formulas are no longer proprietary; analyses can be conducted on commercially available software (e.g., SAS, SPSS)	When analyses are conducted appropriately, findings may be used as one component of NCEES and used for goal-setting at the student-, classroom-, school-, district-, and state-levels Impact will need to be monitored Stakeholder groups will need to be involved in determining cut-points

Model	Face Validity with Stakeholders	Level of Expertise Required to Develop and Use	Ease of Implementation Statewide	Cost of Use/Resource Requirements	Policy Implications
		number of states (e.g., TN, OH, PA) and large districts (e.g., NYC)	through the GM Pilot		for each effectiveness level
Student Fixed Effects Model (SFE)	May be perceived as more reasonable by elementary teachers than by middle school teachers as "student effect" is assumed to be constant (i.e., does not vary across teachers)	Model is relatively less complex but specification still requires expertise Analyses can be conducted by DPI DPI would need to develop reporting strategies	Excludes grade 4 teachers (2 yrs of prior data needed) Allows for statewide (cross-teacher) comparison of teacher effect (school-level variables not required)	No licensing fees or software purchase required Relatively simple, statistically parsimonious and economical	When analyses are conducted appropriately, findings may be used as one component of NCEES and used for goal-setting at the student-, classroom-, school-, district-, and state-levels Impact will need to be monitored Stakeholder groups will need to be involved in determining cut-points for each effectiveness level
Multivariate Response Model (EVAAS-MRM)	May be perceived as stronger because it has gained credibility as the "gold standard" if a state/district can afford the cost	Analyses go through SAS Insitute and DPI uses their results; analytic processes are less transparent but many conveniences as SAS also develops reports	Adjustments to model may allow for statewide adoption, though current model is best implemented at the district level	Proprietary software; license is costly	When analyses are conducted appropriately, findings may be used as one component of NCEES and used for goal-setting at the student-, classroom-, school-, district-, and state-levels Impact will need to be monitored Stakeholder groups will need to be involved in determining cut-points for each effectiveness level Policymakers may feel more confident about use of findings for high stakes purposes if experts from SAS Institute support analyses