Value-Added Modeling Basics

Jennifer Green
Montana State University

Pam Fellers
University of Nebraska-Lincoln
Race to the Top (2009)

• Raise achievement and close gaps
  – Develop common standards and high-quality assessments
  – Build data systems:
    • Measure student growth
    • Inform improvements in instruction
  – Develop & support effective teachers and leaders
  – Turn around lowest-achieving schools
Math & Science Partnerships

• Significant investments in teacher education and development

Does it work?

If so, when? How? For whom? What does effective teaching look like?
Value-Added Modeling

• Dr. William Sanders & Dr. Robert A. McLean
  – Tennessee Value-Added Assessment System (TVAAS)

• Estimate effects of educational factors (P.D., teachers, schools, districts, etc.) on student learning

• Identify characteristics of highly effective programs / teachers
  – Informed improvements in education (McCaffrey et al., 2003)
Outline

• Layered VAMs
• Education Value-Added Assessment System (EVAAS) Model
• Extensions to Professional Development
• Discussion
What is Value-Added?

Test Score

Diff between student score & district avg.

Year

Expected growth for student
District avg.

Ballou, Sanders, & Wright (2004)
What is Value-Added?

Ballou, Sanders, & Wright (2004)
What is a Layered Model?

Test Score

Above district avg. growth

Equal to district avg. growth

Which teacher?

Teacher 1

Teacher 2

District avg.

Year

g

Teacher 1

Teacher 2

g + 1

g + 2

g

Which teacher?
Layered vs. Non-layered Models

• Layered Model:
  – Links past teachers to subsequent student outcomes
    \[
    \text{score}_{g+1} = \text{district mean}_{g+1} + \text{teacher}_1 + \text{error}_{g+1}
    \]
    \[
    \text{score}_{g+2} = \text{district mean}_{g+2} + \text{teacher}_1 + \text{teacher}_2 + \text{error}_{g+2}
    \]

• Non-layered Model:
  – Links current teachers to student outcomes, so ignores effects of instruction in earlier years
    \[
    \text{score}_{g+1} = \text{district mean}_{g+1} + \text{teacher}_1 + \text{error}_{g+1}
    \]
    \[
    \text{score}_{g+2} = \text{district mean}_{g+2} + \text{teacher}_2 + \text{error}_{g+2}
    \]
Linear Mixed Model

\[ y = X\beta + Zu + e \]

Random Teacher Effects:
\[ u \sim N(0, D) \]

\[ \text{Cov}(u, e) = 0 \]

\[ V(y) = ZDZ^T + R \]

Random Error:
\[ e \sim N(0, R) \]
## Example

<table>
<thead>
<tr>
<th>Student</th>
<th>Year</th>
<th>Teacher</th>
<th><strong>Z (Non-layered)</strong></th>
<th><strong>Z (Layered)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Bill</td>
<td>1</td>
<td>A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>B</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sally</td>
<td>1</td>
<td>A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>C</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Example

\[ y_{i1} = \mu_1 + T_A + e_{i1} \]
\[ y_{i2} = \mu_2 + T_A + T_B + e_{i2} \]

- \( y_{i1}, y_{i2} \) = Bill’s scores at time \( g = 1, 2 \)
- \( \mu_1, \mu_2 \) = district-level mean at time \( g = 1, 2 \)
- \( T_A, T_B \) = effects of teachers A, B
- \( e_{i1}, e_{i2} \) = random errors at time \( g = 1, 2 \)

<table>
<thead>
<tr>
<th>Year</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
</tr>
</tbody>
</table>

\( Z \) (Layered)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
EVAAS Model (Sanders et al., 1997)

\[ y = \mu + Zt + e \]

- \( y \) = vector of test scores
- \( \mu \) = vector of means
- \( t \) = random teacher effects, \( t \sim N(0, \sigma_t^2 I_q) \)
- \( e \sim N(0, I_n \otimes \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1,s_k} \\ \sigma_{21} & \sigma_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{s_k,1} & \cdots & \sigma_{s_k,(s_k-1)} & \sigma_{s_k,s_k} \end{bmatrix}) \)
Technical Requirements

• Assessment reflects student achievement with respect to the curriculum objectives

• Assessment must have ability to adequately score all levels of achievement (i.e. free of floor or ceiling effects)

• VAM is valid only to the extent that it has been applied over multiple years of data
  – Quality baseline characterizing starting point
  – Track students over a minimum of three additional years
  – Reliable assessment with equivalent meaning over the modeling period

• Models assume randomization
Estimating Teacher Effects

• Unexplained classroom-level heterogeneity

• Best Linear Unbiased Predictors (BLUPs)

\[ \hat{t} = DZ^T V^{-1} (y - X\hat{\beta}) \]
where \[ \hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} y \]

- “Shrinkage Estimates”
  - Estimates pulled to overall mean, depending on sample size & variability within and across teachers’ classrooms

- Individual teacher effect estimates relative to others in sample
Estimating Teacher Effects

- Standard errors tend to be large
  - Improvement: Yes
  - High-stakes evaluation: Caution

- Could VAM be extended to evaluate program effectiveness?
Professional Development

Theory of action underlying PD programs

Improve/ Foster Teaching

- Content Knowledge
- Pedagogical Knowledge
- Attitudes/Beliefs toward discipline, teaching, learning, etc.

Improve Student Learning
Professional Development

• “Experimental Units” of programs are teachers
  – Teacher level is where to look for impact

• Rationale for VAM
  – Teacher level estimates through longitudinal student learning measures
What is a Program Effect?

Diagram:
- Test Score vs. Year
- Points for Teacher 1: g, g + 1, g + 2
- Points for Teacher 2: g + 1, g + 2

Lines:
- Above district avg. growth
- Equal to district avg. growth
- Program
- Teacher
- District avg.
What is a Program Effect?

• Layered Model with Program Effect

\[
\text{score}_{g+1} = \text{district mean}_{g+1} + \text{teacher}_{1,P} + \text{error}_{g+1}
\]

\[
\text{score}_{g+2} = \text{district mean}_{g+2} + \text{teacher}_{1,P} + \text{teacher}_{2,N} + \text{error}_{g+2}
\]

• Definition?

\[
\text{program effect} = \text{teacher}_{1,P} - \text{teacher}_{1,N}
\]

→ For teachers in the program

– need to know their effect before as well as during
and/or after the program

– need some assurance their effect is stable
What is a Program Effect?

Cohort 1 Student

- Value-added before program
- Above district avg. growth

District Average

g Teacher 1 before PD

g+1

Cohort 2 Student

Value-added after program

- Above district avg. growth

Add’l gain of PD program

Value-added before program

District Average

Teacher 1 with PD

g+1

g+2

Year
## Modified Matrices and Equation

Suppose teacher A was involved in the PD program between Cohort 1 and 2

<table>
<thead>
<tr>
<th>Student</th>
<th>Year</th>
<th>Cohort</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill</td>
<td>1</td>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>B</td>
</tr>
<tr>
<td>Sally</td>
<td>2</td>
<td>2</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>C</td>
</tr>
</tbody>
</table>

### $Z_N$ (No PD)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sally</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

### $Z_P$ (PD)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sally</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[
y = \mu + Z_N t_N + Z_P t_P + e\]
Evaluating Program Effectiveness

- Need counterfactual (e.g., comparison group)

- Need high-quality data
  - Longitudinal (multiple years)
  - Assessment(s) aligned with program goals & able to score all levels of achievement
  - Accurate teacher links

- May not detect significant program effect using “teacher” VAM estimates
  - Use multiple metrics for evaluation, not just student achievement

- Improvement vs. High-Stakes Decisions
Discussion

• How could value-added modeling be used to inform and strengthen teaching?