Estimating Schooling Effects

Daniel Anderson

University of Oregon, EDLD 655

EDLD 655

University of Oregon

Background

School Evaluation

- Two broad reasons for estimating a "schooling effect" on student achievement
  - Research
    - e.g. What makes one school more or less effective than another?
  - School Accountability
- Three broad methods for estimating school effects
  - "Snapshot" estimates
    - e.g. percent proficient
  - Year-to-year gain-scores
    - Estimates the gain in student test scores from one year to the next
  - Longitudinally
    - Uses multiple years of data to fit a growth function to the entirety students' data

Goal

The goal of this project is to determine the most appropriate method for estimating a "school effect" on student achievement.

Why does this matter?

Estimating a schooling effect is critical to both research and accountability.

For **research**, if we fail to adequately measure school effectiveness, then our ability to identify factors that influence school effectiveness are limited.

For **accountability**, if we fail to adequately measure school effectiveness, then schools will be unfairly rewarded or penalized.

<u>The Big Picture</u>

Different methods for estimating school effects can provide ***substantially*** different pictures of school effectiveness (e.g., Zvoch & Stevens, 2008).

Three General Methods

- **"Snapshot" estimates**
  - Used under the *No Child Left Behind Act* (NCLB, 2001)
  - Estimates based off a single test score from a single point in time (e.g., Percent Proficient).
- **Year-to-year gain-scores**
  - Estimates school effects based on students test score gains from one year to the next
  - Many accountability metrics moving this direction (e.g., value-added models)
- **Longitudinal**
  - Uses multiple years of data to fit a growth function to the entirety students' data
  - Most commonly used in research

Contact Information and Acknowledgements

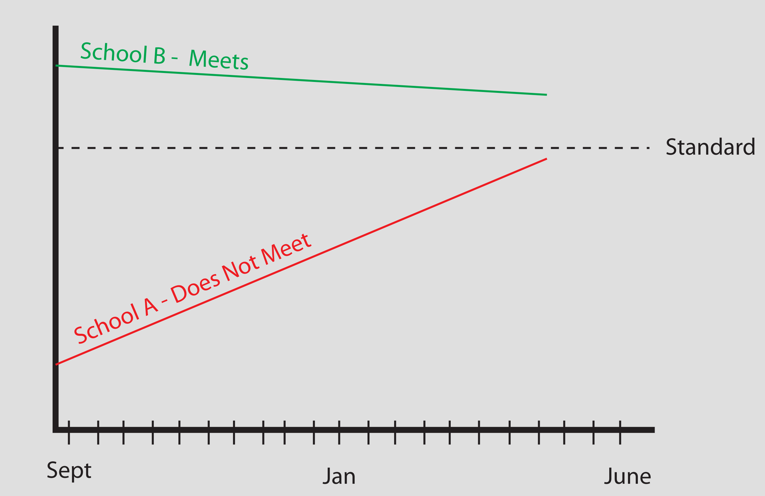
For further information, please contact Daniel Anderson, daniela@uoregon.edu.

Benefits and Challenges For Each Method

Snapshot Methods

Benefits: Easily interpretable – if students don't reach a performance level expectation, the school doesn't either.

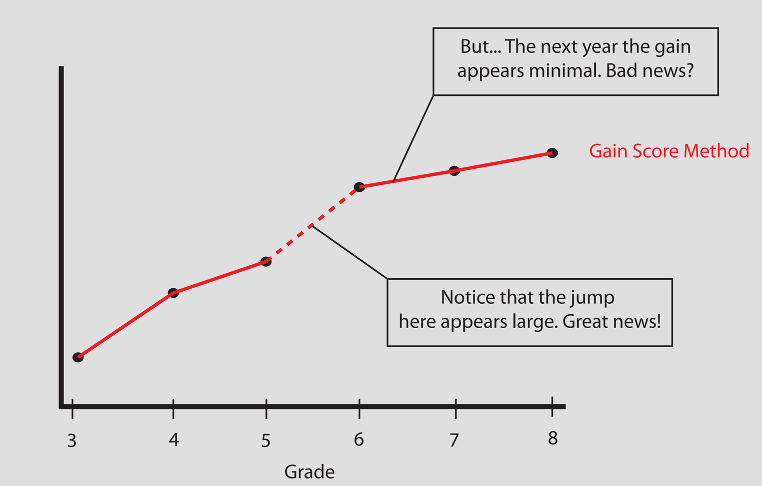
Challenges: Snapshot methods do not account for any student growth that occurs. Estimates have been shown to be biased (Kim & Sunderman, 2005; Zvoch & Stevens, 2008).



Year-to-Year Gain Methods

Benefits: Account for student gains from one year to the next.

Challenges: Not true "growth". Reliability of gain scores are often low (Singer & Willett, 2003). Far more complex computationally, requiring clear communication for stakeholders to understand how schools are being evaluated. Linear (straight line) growth from one year to the next must be assumed.



Longitudinal Methods

Benefits: Provide the most complete account of student learning, which often leads to more stable and reliable school effect estimates.

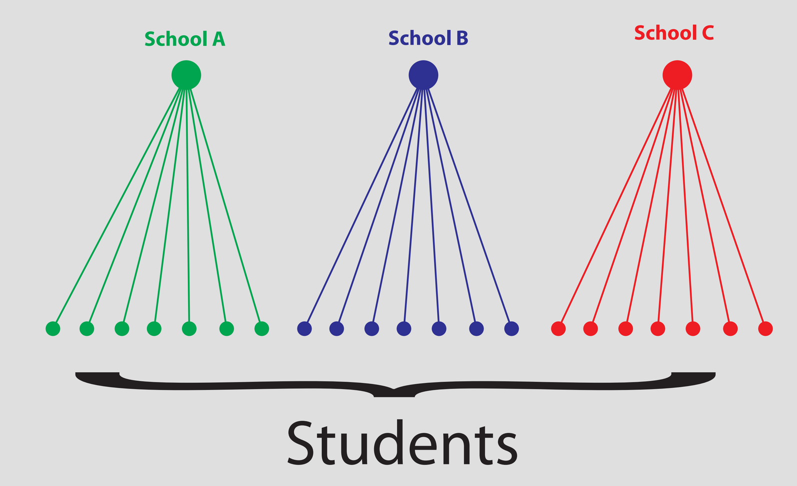
Challenges: Computationally complex. Essentially all the oft-discussed issues related to longitudinal modeling apply (see Singer & Willett, 2003). Handling of student mobility can be particularly tricky, as typical applications require a "pure" nesting structure.

Recommended Method - Longitudinal

Nesting Structures

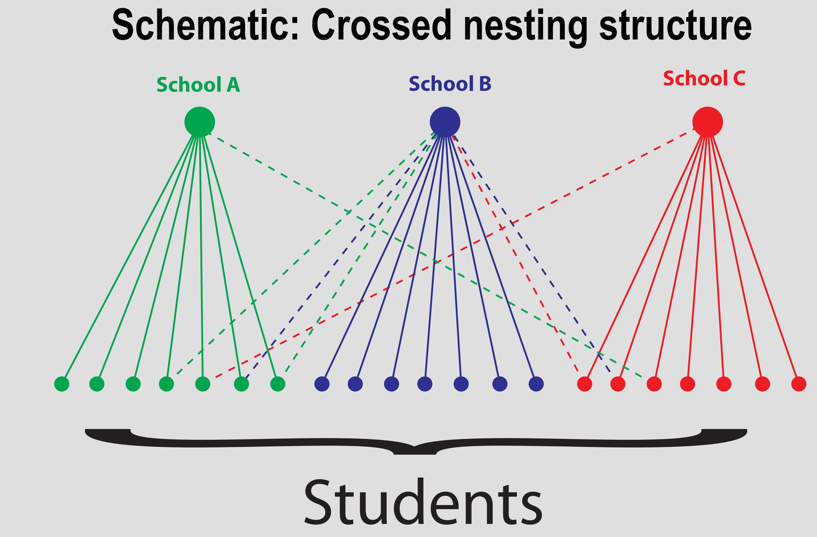
Perhaps the greatest challenge to longitudinal designs is that they typically require a "pure" nesting structure. That is, every student is a member of one, and only one school.

Schematic: Pure nesting structure



Thus, in any case where a student transitions schools, the model breaks down. This, of course, makes the models impractical for accountability purposes given that, in real world data, nearly every student makes a transition at some point (e.g., elementary to middle school).

Schematic: Crossed nesting structure



Dealing with crossed nesting structures

First/Last School Method

Researchers applying the first or last school method to studying school effects treat the entirety of the students' data as being representative of *only* the first or last school attended. This method biases school effect estimates by attributing student data as representative of the school when the student *no longer attends the school* (See Grady & Beretvas, 2010).

Listwise Deletion Method

Researchers applying listwise deletion remove *all* students from the dataset who transition schools *at any point* during the study. This method makes studying any time frame including a middle-school transition impossible. Estimates remain biased due to the final sample being unrepresentative of the full sample (unless mobility is low, See Luo and Kwok, 2012).

Cross-Classified Models

An alternative is to model the observed relationships with students being members of each school. Models are considerably more complex, but provide the best representation of the observed relationships.

Practical Repercussions

Research

The choice of one model for school effectiveness over another can have substantial impacts on the overall "picture" of school effectiveness. In research, snapshot methods are used relatively infrequently. Instead, gain-score or longitudinal models are employed, which each typically produce better overall estimates of school effectiveness (Zvoch & Stevens, 2008). Yet, even when longitudinal models are employed a listwise deletion or first/last school method is generally adopted (Grady & Beretvas, 2010).

Inadequate estimates of school effects in ***research*** *could potentially lead to:*

- School-wide reform efforts being inappropriately labeled as "effective" of "ineffective"
- Schools appearing (incorrectly) to exhibit bias for or against specific student subgroups.
- Resources being inappropriately devoted to studying "outlier" schools that are particularly effective despite circumstances

Accountability

For accountability, snapshot methods have historically been the norm (see NCLB, 2001). However, they have come under considerable criticism given some of the critiques outlined here. Thus, accountability policies have begun to focus more on student learning. To date, however, these have mostly all fallen under the "year-to-year gain-score" method.

Inadequate estimates of school effects in ***accountability policies*** *could potentially lead to:*

•Schools being unfairly labeled as "failing" or "succeeding"
- Schools are then unfairly sanctioned or rewarded
- Public trust in the effectiveness of specific schools may be unduly high or low

•May lead to a "rich get richer while the poor get poorer" dilemma, given that some methods have been shown to be biased against schools serving high proportions of ethnic minorities and students from impoverished backgrounds (Kim & Sunderman, 2005; Zvoch & Stevens, 2008).

References

Grady, M. W., & Beretvas, S. N. (2010). Incorporating student mobility in achievement growth modeling: A cross-classified multiple membership growth curve model. *Multivariate Behavioral Research, 45*, 393-419. doi: 10.1080/00273171.2010.483390

Kim, J., & Sunderman, G. (2005). Measuring academic proficiency under the no child left behind act: Implications for educational equity. *Educational Researcher, 34*, 3-13. doi: 10.3102/0013189X034008003

Luo, W., & Kwok, O. (2012). The consequences of ignoring individuals' mobility in multilevel growth models: A monte carlo study. *Journal of Educational and Behavioral Statistics, 37*, 37-56. doi: 10.3102/1076998610394366

The No Child Left Behind Act, Pub. L. No. 107-110 (2001).

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurence*. New York, NY: Oxford University Press.

Zvoch, K., & Stevens, J. (2008). Measuring and evaluating school performance: An investigation of status and growth-based achievement indicators. *Evaluation Review, 32*, 569-595. doi: 10.1177/0193841X08320398

www.PosterPresentations.com