The Logic and Illogic of Teacher Evaluation

Mark D. Reckase

Some Initial Thoughts

- Student achievement is a result of the interaction of the student and the educational environment including each teacher.
- Teachers are responsible for creating the situation that enables learning and facilitating the learning process.
- It seems reasonable to evaluate teachers at least partially by the achievement of their students.
- But, a teacher can not make a student learn. The characteristics of the students and the educational setting must be taken into account.
- Everything here reflects my views, not those of the MCEE.

Organization of Presentation

- Review some of the details of the legislation calling for the Educator Evaluation System.
- Explain some of the procedures for estimating student growth in achievement and the value-added attributed to the teacher.
- Discuss the credibility of student growth and value-added estimates.
- Describe some of the procedures for teacher observation.
- Consider the implications of the requirements of the Educator Evaluation System.

Act No. 102
Public Acts of 2011
Approved by the Governor
July 19, 2011
Filed with the Secretary of State
July 19, 2011
EFFECTIVE DATE: July 19, 2011

STATE OF MICHIGAN
96TH LEGISLATURE
REGULAR SESSION OF 2011

Introduced by Rep. O'Brien

ENROLLED HOUSE BILL No. 4627

AN ACT to amend 1976 PA 451, entitled "An act to provide a system of public instruction and elementary and secondary schools; to revise, consolidate, and clarify the laws relating to elementary and secondary education; to provide for the organization, regulation, and maintenance of schools, school districts, public school academies, intermediate school districts, and other public school entities; to prescribe rights, powers, duties, and privileges of schools, school districts, public school academies, intermediate school districts, and other public school entities; to provide for the regulation of school teachers and certain other school employees; to provide for school elections and to prescribe powers and duties with respect thereto; to provide for the levy and collection of taxes; to provide for the borrowing of money and issuance of bonds and other evidences of indebtedness; to establish a fund and provide for expenditures from that fund; to provide for and prescribe the powers and duties of certain state departments, the state board of education, and certain other boards and officials; to provide for licensure of boarding schools; to prescribe penalties; and to repeal acts and parts of acts," by amending section 1249 (MCL 380.1249), as amended by 2010 PA 336, and by adding sections 1248 and 1249a.

Requirements of Legislation

- The board … shall adopt and implement for all teachers and school administrators a rigorous, transparent, and fair performance evaluation system that does all of the following:
  - Evaluates the teacher's or school administrator's job performance at least annually while providing timely and constructive feedback.
  - Establishes clear approaches to measuring student growth and provides teachers and school administrators with relevant data on student growth.
  - Evaluates a teacher's or school administrator's job performance, using multiple rating categories that take into account data on student growth as a significant factor.

Requirements of Legislation (2)

- For these purposes, student growth shall be measured by national, state, or local assessments and other objective criteria.
- it rates teachers as **highly effective, effective, minimally effective, or ineffective**.

- Uses the evaluations, at a minimum, to inform decisions regarding all of the following:
  - The effectiveness of teachers and school administrators, ensuring that they are given ample opportunities for improvement.
  - Promotion, retention, and development of teachers and school administrators, including providing relevant coaching, instruction support, or professional development.
  - Whether to grant tenure or full certification, or both, to teachers and school administrators using rigorous standards and streamlined, transparent, and fair procedures.

Legislative Requirements (3)

- Removing ineffective tenured and untenured teachers and school administrators after they have had ample opportunities to improve, and ensuring that these decisions are made using rigorous standards and streamlined, transparent, and fair procedures.

Specific Requirements

- For the annual year-end evaluation
  - for the 2013-2014 school year, at least 25% of the annual year-end evaluation shall be based on student growth and assessment data.
  - For the annual year-end evaluation for the 2014-2015 school year, at least 40% of the annual year-end evaluation shall be based on student growth and assessment data.
  - Beginning with the annual year-end evaluation for the 2015-2016 school year, at least 50% of the annual year-end evaluation shall be based on student growth and assessment data.
- The performance evaluation system shall include classroom observations to assist in the performance evaluations.
  - A classroom observation shall include a review of the teacher's lesson plan and the state curriculum standard being used in the lesson and a review of pupil engagement in the lesson.
  - A classroom observation does not have to be for an entire class period.

Specific Requirements (2)

- All student growth and assessment data shall be measured using the student growth assessment tool that is required under legislation enacted by the legislature under subsection (6) after review of the recommendations contained in the report of the governor's council on educator effectiveness submitted under subsection (5).

Details – Student Growth

- If there are student growth and assessment data available for a teacher for at least 3 school years, the annual year-end evaluation shall be based on the student growth and assessment data for the most recent 3-consecutive-schoolyear period.
- If there are not student growth and assessment data available for a teacher for at least 3 school years, the annual year-end evaluation shall be based on all student growth and assessment data that are available for the teacher.

Details – Student Growth

- Is a value-added model that takes into account student achievement and assessment data, and is based on an assessment tool that has been determined to be reliable and valid for the purposes of measuring value-added data.
- In addition to measuring student growth in the core subject areas of mathematics, science, English language arts, and social science, will measure student growth in other subject areas.
- Complies with all current state and federal law for students with a disability.
- Has at least a pre- and post-test.
- Is able to be used for pupils of all achievement levels.

Evaluation Tool for Teachers

- May include, but is not limited to, instructional leadership abilities, teacher and pupil attendance, professional contributions, training, progress report achievement, school improvement plan progress, peer input, and pupil and parent feedback.
- Will allow all special education teachers to be rated.

Evaluation System for Administrators

- Also based on student growth with the same proportions as the teachers.
- Proficiency in using the teacher evaluation tool.
- The progress made by the school or school district in meeting the goals set forth in the school's school improvement plan or the school district's school improvement plans.

Thoughts about the Legislative Requirements

The Allure of Value-added Models

- It is not surprising that parents and school policy makers want to know how much students gain from the educational process.
- On average, how much do students learn during a typical academic year?
- Do low performing or high performing students learn more during the year?

The Allure of Value-added Models

- Do students at some schools learn more than students at other schools?
- Do students with some teachers learn more than those with other teachers?
- How can we estimate these differences in a way that is fair and accurate for all of them involved?

What Is a Value-added Model?

- One way of estimating value-added is based on a model of achievement growth that is called an educational production function. There are other methods as well.
- An educational production function shows the relationship between end of year achievement and variables that influence that achievement.

The Educational Production Function

- $A_{it} = f_t\ (X_{it},\ldots,X_{i0},\ E_{it},\ldots,\ E_{io},\ c_i)$
- Achievement of student *i* at time *t* (A_{it}) is a function at time *t* of
  - Student and family inputs (*Xs*)
  - School inputs including teacher effects (*Es*)
  - A general student effect (c_i).
- Does this general expression seems reasonable?

Education Production Function

- Although the general function can have any form, most of the models used in practical settings are based on linear relationships.
  - This is typical – most scientific models begin as linear models.
  - Most statistical models used to analyze educational data are linear models.

A Linear Model

- $A_{it} = \alpha_t + E_{it}\beta_0 + E_{i,t-1}\beta_1 + \ldots + E_{i0}\beta_t + X_{it}\gamma_0 + X_{i,t-1}\gamma_1 + \ldots + X_{i0}\gamma_t + \eta_t c_i + \mu_{it}$

- In this equation, the *E*, *X,* and *c* variables are as defined earlier and the multipliers are regression weights.
  - α is an adjustment if the achievement measures from grade to grade are not on the same scale.
  - μ is all of the unmeasured influences.

A Simpler Linear Model

- Achievement in year *t* is assumed to be a function of
  - Achievement in year *t* – 1,
  - The student general effect, *c*,
  - The teacher effect, *E,*
  - And the unmeasured influences
- $A_{it} = A_{i,t-1} + c_i + E_{teach} + \mu_i$
- All of these are assumed to be measured without error.

A More Complex Model

- $A_{it} = LA_{i,t-1} - d + c_i + E_{teach} + \mu_i$
- The multiplier *L* and the term *d* define linear decay in what is learned. In other words, students forget part of what they learned in the previous year.
  - The magnitude of *L* (between 0 and 1) reduces the variance of the previous years score.
- The variables A_{i0} and c_i are allowed to be correlated.

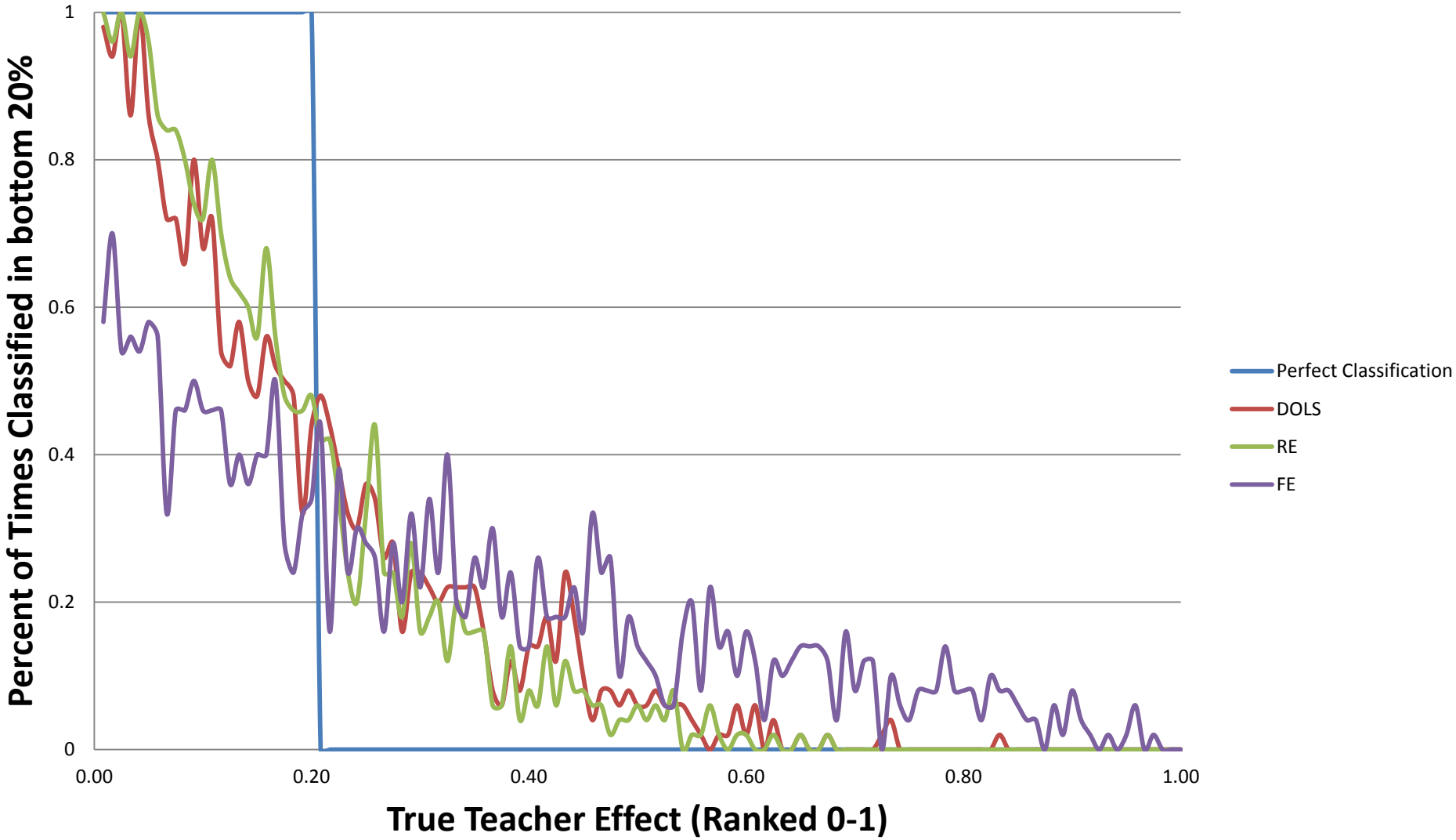
Under Random Assignment Most Approaches Work Fairly Well

- As long as students are randomly assigned to teachers, most of the estimators achieve a fair amount of accuracy in estimating and ranking teacher effects
- This is particularly true if teachers are randomly distributed across schools
- Grouping students (tracking) in itself does not cause problems

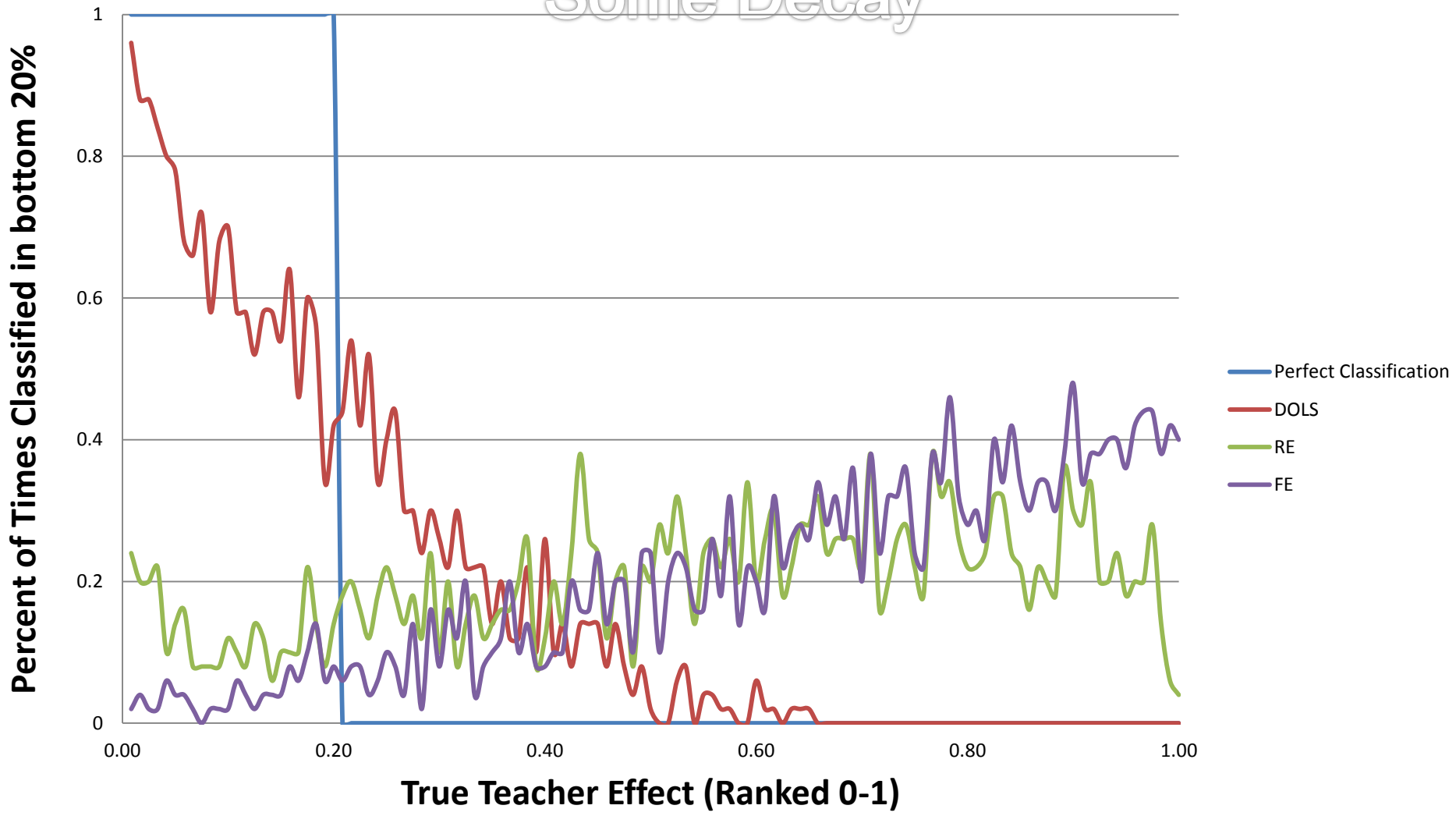
A Sample of Some Findings

Scenario	Summary Statistic	DOLS	RE	FE
Random grouping Random assignment No Decay	Rank correlation	.84	.88	.49
	Proportion misclassified as below average	.17	.14	.50
Grouping by prior test scores Best students to best teachers Some Decay	Rank correlation	.82	-.09	-.35
	Proportion misclassified as below average	.20	.52	.74

Identifying the "Worst" Teachers – Best Case Scenario
Random Assignment



Identifying the "Worst" Teachers – More Difficult Scenario
Best Students Assigned to Best Teachers—Some Decay



Classroom Observation

- There are numerous observational methods
  - Charlotte Danielson's Framework for Teaching
  - Marzano Teacher Evaluation Model
  - The Thoughtful Classroom
  - 5 Dimensions of Teaching and Learning
- All require training
- All recommend multiple observations
- To give good information about teachers, need
  - Representative sampling of instructional time
  - Sufficient observations to get reliable estimates of the subscores from the observations.

Classroom Observations

- The training requirements are extensive, but they are necessary to have comparable ratings from different individuals.
- The multiple observations requirement create challenges for scheduling and finding sufficient staff time.
- Different classroom/instructional models (i.e., team teaching) add complications to the process.
- If multiple methods are allowed, some procedure will be needed to make the results comparable.
- Some credible method must be developed to combine observational data with measures of growth.

Reactions

- There are hidden complexities to the process
  - Some teachers have large numbers of students and data from multiple-years
  - Some teachers have results on their students from multiple tests
  - Some teachers teach more than one course (e.g., algebra and geometry, French and history).
  - Contact hours with students vary across teachers – student attendance differs for different schools.
  - There are many more.
- A final approach must determine how to deal with all of these complexities.

Reactions

- The value-added measures by themselves are not sufficiently precise to make high stakes decisions.
- As with other test-based measures, more than one information source is needed to get accurate estimates.
- In this case, the value-added measures will be augmented with classroom observations and other indicators of good practice.
- A pilot test was done to obtain information about the possible procedures.
- Ultimately, high stakes decisions should not be made using information from one year – patterns over years need to be documented.

The Logic

- The major job of a teacher is to facilitate student learning, therefore it makes sense to evaluate teachers based on how much students learn.
- Multiple measures are needed to get good information about the performance of the teachers.
- The methodology for evaluating teachers must take into account the students assigned to the teachers and the teaching environment.

The Illogic

- The legislation is very specific about some things that are very technical without considering the nuances of the technical decisions.
  - The weighting of components of the system.
  - The selection of growth measures.
  - The timing of the procedures.
- Other parts of the legislation are not very specific.
  - The definition of a teacher is not provided
  - The challenges of evaluating teachers in non-tested subjects are not considered.
  - Data systems and implementation issues are not considered.