1. Questions(s) or issue(s) for dialogue at Learning Network Conference session:

Do MSPs improve student outcomes? As MSPs move from small, single school or single district experiments to large scale partnerships, the role of evaluation in assessing what we can learn from the interventions escalates in importance. This paper aims to inform new evaluation projects by describing a model for assessing the effect of teachers’ participation in a MSP on their students’ outcomes. The paper focuses on the data requirements of the model and the data challenges for a large scale evaluation project that is being conducted in central Appalachia (the Appalachian Math and Science Partnership (AMSP)).

2. Context of the work within the STEM education literature and within your MSP project:

The passage of No Child Left Behind (NCLB) has had a perhaps unintended consequence of greatly expanding the teacher quality literature. Because of accountability requirements, states are not only tracking student progress in the STEM areas but in education more generally. Tracking student progress means that states must develop testing that is applied to all students of at various grade levels.

Using the student testing data, researchers have increasingly focused on the educational inputs that contribute to successful outcomes. The one policy variable consistently found to be significant in these studies in teacher quality. As a consequence, research efforts have moved toward identifying observable characteristics of teachers, including pre-service and in-service training that improves teacher quality. This paper and this RETA project build upon this growing literature in the education policy world that examines the quality of teachers by looking at the effect of the MSP training on student test scores over time.1 In particular, the training received by teachers in the AMSP is viewed as in-service training directed toward improving teacher quality and ultimately improving student scores.

There are alternative possible measures of the success of MSP-type teacher training. Some scholarly work has looked at the partnerships themselves and their effects on permanently changing school cultures, and others focus on the teachers. The latter literature typically looks at

---

1 See “Special Issue: Key Issues in Value-Added Modeling,” Education Finance and Policy (2009), Issue 4, Fall, for a series of articles on the econometric and psychometric challenges in this type of modeling.
pre- and post-tests of teacher content or pedagogical knowledge. But ultimately, the purpose of STEM interventions generally, and the MSPs in particular, is to increase student learning through enhanced teacher quality.

3. Claim(s) or hypothesis(es) examined in the work (anticipating that veteran projects will have claims, newer projects will have hypotheses):

We hypothesize that teacher participation in an AMSP (or MSP more generally) activity leads to higher student outcomes for that teacher’s students relative to students of teachers who have not participated. We also propose to examine the degree of intervention necessary to generate the hypothesized positive effects.

While enhanced student outcomes are the goal of all MSPs, one advantage of looking at the AMSP participation is the length of time the training program has been in place (it was one of the original MSPs) and the scale of the project. Because of the long length of time, we can test whether there are delayed effects of the teacher training on student outcomes. The project’s implementation across 54 diverse school districts in 4 states also allows us to control for many factors that may influence outcomes that cannot be garnered from a few, similar school districts. For example, the socioeconomics aspects of the student body, the quality of teacher pre-service training, and the culture of the schools involved in AMSP vary greatly and can be accounted for in this evaluation. Overall, this means that other factors that may complicate attribution to the in-service MSP training can be controlled with this study.

4. Evaluation and/or research design, data collection and analysis:

To assess whether AMSP professional development activities for teachers have had an effect on student outcomes, we estimate a standard education production function or value-added model. These models estimate test scores, or student achievement, over some given time period as the dependent variable. A student’s score at any given time period is expected to be influenced by the student’s own past scores, his or her own attributes such as race and gender, his or her family characteristics such as education of the parents and wealth of the family, the quality of the teacher in that time period, and the characteristics of the school. Variables to represent these factors are typically included as independent variables in such models. For our purposes, it is important to include a variable(s) to measure the AMSP intervention. In other words, we are most interested in assessing whether a teacher’s participation in AMSP and the extent of that participation influenced her student’s achievement while controlling for other factors that also are expected to influence student achievement. Thus, we estimate the following for students in a given year and across multiple years:

---


Student Outcomes = f(Student Characteristics, Student Past Scores, Family Attributes of the Student, School Characteristics, Teacher Characteristics, Teacher AMSP Participation).

Furthermore, because teachers may not randomly select into professional development activities including the MSPs, the model must correct for possible biases regarding which teachers are participating. There are alternative ways to correct for these biases and the paper discusses each briefly.

Note that this equation requires that student scores for a given cohort of students on a particular subject area or grade level must be tracked over time. Additionally, student characteristics such as the student’s gender, race, and socioeconomic status must be available to the researcher. And, often most challenging of all, the individual student must be linked to the specific teacher that taught the student in the particular subject area as well as linked to the characteristics of the teacher. All of this requires multiple periods of observation so that past scores as well as present scores of students can be captured. We also want to observe the student with multiple teachers and the teachers with multiple students.

New federal policies encourage states to collect individual student and individual teacher data and make the data available for research purposes. To date, the AMSP states do not capture these data at a centralized level (and only a few states do). This paper describes explicitly the process of collecting data from a variety of decentralized sources that enable this evaluation to take place. The data collection process is complicated by the fact that most of the data are confidential. Beginning with the step of soliciting approval to retrieve school roster files that document which students are enrolled in the class of which teacher to the final step of matching the various necessary pieces of data, we describe the minimally necessary data requirements for evaluating the effects of MSPs on student outcomes using a general research design as described above.

5. Key insights (retrospective for veteran projects, prospective for newer projects) that have value for the Learning Network:

By sharing the insights we have gained about data collection and evaluation design, new projects can benefit both from our early mistakes and from our successes. Our RETA project is in the final stages of data collection and early stage of data analysis. It is an excellent point to share insights with other MSPs.

This type of evaluation of the AMSP was not an integral part of the initial phase of the MSP. As a result, much of the data are being collected retrospectively and this in itself presents special challenges. But this also suggests that other projects that have not incorporated these types of evaluation models into their initial project designs may be able to consider altering their types of analyses and data collections to generate similar evaluation designs. The project insights regarding both the model and the data requirements for such modeling should be equally valuable to new projects that wish to adopt this type of evaluation design.