PostDoc Journal
Vol. 1 No. 5, May 2013

Journal of Postdoctoral Research
www.postdocjournal.com

Teacher Evaluation and Value-Added: Do Different Models Give Us the Same Answer?

Guodong Liang[1,*]

[1]*Community Training & Assistance Center, 30 Winter St, Boston, MA, 02108, USA*

Abstract This research highlight reviewed the study of Goldhaber, Goldschmidt, and Tseng (2013) on the estimates of teacher effectiveness and the specification of value-added models.

Key words: Teacher training, teacher assistance, teacher evaluation, SLOs, VAMs

Teacher evaluation plays a key role in enhancing teachers' effectiveness, promoting professional learning, and making personnel decisions in the United States and around the world (e.g., Liang, 2013). In addition to Student Learning Objectives (SLOs), an effective tool in evaluating teachers who do not teach subjects or grades that are assessed with state assessments (Slotnik & Smith, 2004, 2013), in recent years, value-added models (VAMs) have become a leading candidate for estimating individual teachers' contribution to student achievement and growth on state standardized tests. However, there are concerns on such key issues as the reliability and validity of VAMs. Empirical studies have shown that teacher ratings with VAMs are highly instable and can vary significantly across time, classes, and even tests (Lockwood et al., 2007; Loeb & Candelaria, 2013; Newton, Darling-Hammond, Haertel, & Thomas, 2010).

Existing research on VAMs focuses on elementary and middle school teachers because their students are often tested consecutively across the years/grades. A standard VAM framework often regresses student achievement in one year/grade against the achievement in the prior year/grade, controlling for a set of school and/or teacher variables and individual covariates. Because the No Child Left Behind (NCLB) act only requires one high-school grade to be tested, there is often a lack of comparable tests that measure student achievement in a prior grade in high school classes. It is therefore challenging to estimate teacher effects at the high-school level with VAMs.

In a recent study published in the *Educational Evaluation and Policy Analysis* (Goldhaber, Goldschmidt, & Tseng, 2013), the researchers compared teacher effectiveness estimates derived from traditional VAMs with lagged scores using pretests and posttests in single subject areas to those from VAMs with a cross-subject student fixed-effects approach. Their study utilized a unique dataset collected by ACT which included 23 schools, 205 teachers, and 8,002 students in grades nine through 12. They found that model specification influences both the estimated impact of teacher quality on student achievement and estimates of individual teacher effectiveness. In particular, teacher effects identified using the within-student cross-subject variation approach are significantly smaller than those generated from the traditional lagged score VAMs.

Although this study cannot answer the critical question of what model specification is likely to provide the most accurate information of true teacher effectiveness, Goldhaber and colleagues (2013) provides strong evidence that the estimated effect size of teacher quality and individual teacher performance depends on model specifications. Given the national enthusiasm for evaluating teachers using student test scores, this study highlights the importance and urgency of more research on the assumptions about what drives

*gliang@ctacusa.com

student learning and VAM specifications at the high school level.

A teacher's value-added score is likely to vary across time, subject, student population, and model specifications (e.g., Goldhaber & Theobald, 2013; Loeb & Candelaria, 2013), however, we know more about VAMs than other commonly used measures of teacher performance such as principal evaluations and classroom observations, and results from VAMs are positively related to those from other approaches (Harris, 2013). When multiple data sources, models, and performance measures across multiple years are used, we are likely to achieve more reliable, valid, and stable estimates of teacher effectiveness (Harris, 2013; Liang, 2013; Loeb & Candelaria, 2013).

Reference

Goldhaber, D. D., Goldschmidt, P., & Tseng, F. (2013). Teacher value-added at the high-school level: Different models, different answers? *Educational Evaluation and Policy Analysis, 35*, 220-236.

Goldhaber, D. D., & Theobald, R. (2013). *Do different value-added models tells us the same things?* Stanford, CA: Carnegie Knowledge Network.

Harris, D. (2013). *How do value-added indicators compare to other measures of teacher effectiveness?* Stanford, CA: Carnegie Knowledge Network.

Liang, G. (2013). Teacher evaluation policies in the United States: Implementation and impact on constructivist instruction. In M. Akiba (Ed.), *Teacher reforms around the world: Implementations and outcomes*. Bingley, UK: Emerald Publishing.

Lockwood, J. R., McCaffrey, D. F., Hamilton, L. S., Stecher, B., Le, V.-N., & Martinez, J. F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement, 44*(1), 47-67.

Loeb, S., & Candelaria, C. (2013). *How stable are value-added estimates across years, subjects and student groups*. Stanford, CA: Carnegie Knowledge Network.

Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives*, *18*. Retrieved April 14, 2011, from http://epaa.asu.edu/ojs/article/view/810

Slotnik, W. J., & Smith, M. (2004). *Catalyst for change: Pay for performance in Denver*. Boston, MA: Community Training and Assistance Center.

Slotnik, W. J., & Smith, M. (2013). *It's more than money: Teacher incentive fund–Leadership for educators' advanced performance Charlotte-Mecklenburg schools*. Boston, MA: Community Training and Assistance Center.