

**Are School Factors Important for Measuring Teacher Effectiveness? A Multilevel
Technique to Predict Student Gains Through a Value-Added Approach**

Bidya Raj Subedi, Ph.D.

School District of Palm Beach County

Bonnie Swan, Ph.D.

University of Central Florida

Michael C. Hynes, Ph.D.

University of Central Florida

Paper presented at the annual meeting of American Educational Research Association,
Denver, Colorado, April 30 - May 4, 2010.

Abstract

This paper investigates the effect of teacher quality, represented by teacher level characteristics, on mathematics gain scores employing multilevel (or three-level hierarchical linear model, HLM, based) unconditional and conditional value-added model (VAM). We found the significant effects of teacher's mathematics content certification, teacher experience, and the interaction effects of mathematics content certification with student level predictors. Although school poverty significantly predicted gain scores, the school effect for predicting gain scores was negligible.

Key words: Teacher effects, school effects, value-added model, hierarchical modelling.

Introduction

Student achievement gain can be predicted due to individual predictors at student, teacher, and school levels. This predictive model generates a multilevel research design, known as hierarchical linear model (HLM), which also allows the cross-level interactions among the predictors at all three levels. This study aims to predict students' mathematics gain scores due to key predictors at student, teacher, and school levels, incorporated in level-1, level-2, and level-3 models respectively, including cross-level interaction terms. We also report the variance explained and effect sizes at school and teacher levels for measuring teacher effectiveness. Specifically, we measure the magnitude of teacher quality, represented by teacher effectiveness, which is based on three teacher level factors: teacher content-area certification, teacher experience and teacher's advanced mathematics or mathematics education degree.

Recent studies address the relationship between student achievement as well as achievement gains and the factors at student, teacher, and school levels mainly by employing sophisticated statistical techniques and use of the large data sets. For example, in a review of multi-level studies relating to teacher quality and student achievement, Scheerens and Bosker

(1997) found that the differences in student achievement are associated with school (20%) and classroom/teacher level factors (20%), with the remaining difference (60%) at the student level factors (such as socioeconomic status and prior achievement). Rowan, Correnti, and Miller (2002) employed three-level HLM using value-added approach to predict mathematics and reading achievement and annual gains incorporating student, teacher, and school level predictors respectively in level-1, level-2 and level-3 models. Rowan et al. allowed variance decomposition among students, classrooms and schools in order to measure teacher effectiveness where they mention that the purpose of value-added models is to estimate the proportions of variance in changes in student achievement lying among classrooms, after controlling for the effects of other confounding variables. Other powerful value-added models (see Jordan, Mendro, and Weerasinghe, 1997; Sanders and Rivers, 1996), that track students' gains over more than one year, have brought about a rethinking among researchers regarding the relative importance of the role of the teacher. Sanders and Rivers' (1996) ground-breaking Tennessee value-added study showed that fifth grade mathematics students matched in performance assigned to ineffective teachers for three years performed dramatically worse (separated by 50 percentile points on comparable assessments) than children assigned to more effective teachers. Similarly, Jordan et al. (1997), who isolated the effects of Texas teachers on student achievement, found differences of 34 percentile points in reading and 49 percentile points in mathematics achievement, when comparing students assigned to ineffective teachers for three consecutive years to students assigned to three years with effective teachers (defined by how much their students improved).

In the context of the mandates and philosophies of the No Child Left Behind (NCLB) Act in United States of America (U.S.A.), much of what is driving educational reform centers on the premise that teachers matter. For example, by the end of the 2005–2006 school year, states were required, for the first time, to have data collection and reporting mechanisms in place to ensure the

ability to publish reports disclosing whether they meet the goal of ensuring all teachers are “highly qualified.” Meeting these standards basically means that teachers must (a) hold an acceptable bachelor’s or higher degree, (b) have state licensure or certification, and (c) demonstrate subject competency of the subject(s) at the grade level(s) taught.

Past research has shown the need for qualified teachers is particularly great in lower-performing schools with higher numbers of low-income and minority students (see Allen, 2005; Betts, Rueben, and Danenberg, 2000; Hanushek et al., 2004; Lankford, Loeb, and Wyckoff, 2002; Sanders, and Rivers, 1996; U.S. DOE, 2005); and the problem is even more pronounced in middle schools (see Jerald and Ingersoll, 2002). The Government Accountability Office (GAO, 2006) projects 20% of middle schools will have a difficult time meeting NCLB provisions for “highly qualified” teachers.

Evidence is mounting that better teachers can and do make a difference in student achievement (Haycock, 1998; Jordan et al., 1997; Sanders and Rivers, 1996). Still, substantial disagreement exists among researchers as to which teacher qualifications make a difference (Greenberg, Rhodes, Ye, and Stancavage, 2004), and little has been explored on this topic specific to the middle school classrooms. Further, Rice (2003) found a serious gap in the knowledge base that still needs to be explored regarding middle schools (and elementary schools) teachers’ effectiveness that is used to guide important teacher policy decisions. Her award winning review examines the impact of teacher characteristics on teacher effectiveness. In a study related to eighth grade students’ mathematics achievement using 1996 National Assessment of Educational Progress (NAEP) data, Wenglinsky (2002) found that the effects of classroom practices, when added to those of other teacher characteristics, are comparable in size to those of student backgrounds, suggesting that teachers can contribute as much to student learning in mathematics as the students themselves. Through a research on teacher qualification, Croninger, Rice, Rathbun,

and Nishio (2007) found potential contextual effects of teachers' qualifications on student achievement, with first graders demonstrating higher levels of reading and mathematics achievement in schools where teachers report higher levels of coursework emphasis in these subject areas.

Darling-Hammond, Holtzman, Gatlin, and Heilig (2005) found that certified teachers consistently produce stronger student achievement gains than do uncertified teachers, and controlling for teacher experience, degrees, and student characteristics, uncertified teachers are less effective than certified teachers. Darling-Hammond (2000) found that measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics, both before and after controlling for student poverty and language status.

Decker, Mayer, and Glazerman (2004) found that teachers recruited through Teach for America (TFA) are significantly more effective than both uncertified and certified teachers at mathematics instruction and statistically indistinguishable in reading instruction. However, Kane, Rockoff, Staiger (2006) found no difference between teaching fellows and certified teachers or between uncertified and certified teachers in their impact on mathematics achievement. Subject content-area certification has a major role in significantly impacting student achievement. For example, Goldhaber and Brewer (1998) found that mathematics teachers who have a standard certification have a statistically significant positive impact on student test scores relative to teachers who either hold private school certification or are not certified in their subject area.

Relevant studies address the effect of teacher degree and experience on student mathematics achievement (see Ballou and Podgursky, 2000; Darling-Hammond, 2000; Goldhaber, and Brewer, 1998; Howley, 1996; Lippman, Burns, and McArthur, 1996; Monk, 1994; Rice, 2003; Swan, 2006). Subedi, Swan, and Hynes (2005) previously found far fewer teachers with

advanced degrees taught in schools with high percentages of students qualifying for free and reduced lunch versus the number of those in those in wealthier schools.

Researchers, in past, have used student level predictors in multilevel model by incorporating students' prior achievement and socioeconomic background in the model to predict mathematics and reading achievement (see Rowan et al., 2002; Scheerens and Bosker, 1997) including cross-level interactions with predictors at higher level to measure school effects (see Subedi, 2007; Pituch, 1999). Through value-added model to measure teacher effects on annual gains in student achievement, Rowan et al. (2002) used students' prior achievement, socioeconomic status, and school poverty to predict students' gain scores employing three-level HLM. Tobe (2009) mentions that the differences between teachers can be quantified as "teacher effects" using value-added models.

Methods

Design of the Study

This study explores the individual effects of teacher level predictors: mathematics content-area certification, advanced mathematics or mathematics education degree, and experience. Subedi, Swan, and Hynes (2009) conducted similar study using two-level HLM analysis. This study extends their work using three-level HLM through value-added model in order to measure teacher effectiveness. Since the students were not allocated randomly within teachers' classrooms, and student, teacher, and school level predictors incorporated in three separate models provide better estimates of variance and predictors' effects, our best choice of statistical design to measure the teacher effects involves selecting a multilevel or an HLM technique (see Goldstein, 1995; Raudenbush and Bryk, 2002; Subedi, 2005).

The following research questions are explored through this study:

1. What are the significant predictors at student, teacher, and school levels, including cross-level interaction terms, for predicting students' gain scores using conditional VAM?
2. What are the proportions of variance explained and effect sizes (for measuring teacher effectiveness) at teacher level and at school level for unconditional model and conditional VAM ?

Data and Variables

This study used 6,184 students and 253 mathematics teachers from all middle schools in the Orange County Public Schools (OCPS), which is one of the largest urban school districts in Florida and the nation. To be more specific, the OCPS was the twelfth largest of more than a total of 16,000 districts in the United States at the time of data collection.

Outcome variable. We used grades 6-8 mathematics gain scores based on NRT-NCE (Norm Referenced Test-Normal Curve Equivalent) portion of the FCAT test scores for 2005 as an outcome variable, and pretest scores for 2004 as one of the predictors at student level. The NCE scores (for 2005 and 2004) ranged from 1 to 99 and gain scores (i.e., the difference between the scores of 2005 and 2004) ranged from -31.4 to 45.

Student level predictors. The student level predictors used in this study were pretest scores (i.e., NCE scores for 2004) and student socioeconomic status (SES). Student socioeconomic status (SES) was coded 1 for participation and 0 for non-participation in the free and reduced lunch program.

Teacher level predictors. Teacher's content-area certification, a dichotomous predictor, is coded as 1 (indicating holding a mathematics content-area certificate by a teacher for Mathematics 5-9 or 6-12) and 0 (indicating not holding such certificate). The advanced mathematics degree, another dichotomous predictor, is also coded as 1 (indicating holding an advanced degree in

mathematics or mathematics education by a teacher) and 0 (indicating not holding such a degree). Note that advanced degree is defined as master or higher level degree. Teacher experience, a continuous predictor, was measured in number of years the teacher taught. This variable ranged from 0 to 37.

School level predictors. School poverty is defined as the percent of free and reduced lunch students in each school, and teachers' school mean experience is defined as the average number of years taught by middle school teachers in a given school. Further, percent of advanced mathematics degree is defined as the percent of teachers in a particular school with advanced mathematics degree.

Model Development

This study employed a three-level HLM where student, teacher, and school data are incorporated in level-1, level-2, and level-3 models respectively to predict students' gain scores. Pretest scores and SES are used as student level predictors at level-1 model. Content-area certification in mathematics, experience, and advanced degree in mathematics or mathematics education are included as teacher level predictors at level-2 model. School poverty and teachers' school mean experience are used as school level predictors at level-3 model.

First, level-1, level-2, and level-3 unconditional models, which did not include any predictors at any level, were developed. The proportion of variance explained and the effect sizes were calculated at teacher and school level models in order to answer research question 2.

In order to predict mathematics gain scores, the unconditional models at level-1, level-2, and level-3 can be developed as follows.

$$\text{Level- 1: } (\text{MATHGAIN})_{ijk} = \pi_{0jk} + e_{ijk} \quad (1)$$

$$\text{Level- 2: } \pi_{0jk} = \beta_{00k} + r_{0jk} \quad (2)$$

$$\text{Level- 3: } \beta_{00k} = \gamma_{000} + u_{00k} \quad (3)$$

In above equations, π_{0jk} , β_{00k} , and γ_{000} are the intercepts and e_{ijk} , r_{0jk} , and u_{0jk} are the error terms at student, teacher, and school level models, respectively. We want to estimate r_{0jk} and u_{0jk} to find the proportion of explained variances and effect sizes based on these variances at teacher and school levels, respectively in order to answer the research question 2.

Since the purpose of value-added model (VAM) is to estimate the proportions of variance in student gain scores lying among teachers after including important predictors in level-1, level-2, and level-3 conditional models, we successively developed such models. The level-2 and level-3 variance terms were deleted from these model if either they were not significant or did not explain more variance in student gain scores after including the error terms in the model. Subedi (2005) suggested the formulation of level-2, and level-3 conditional models only after the evidence of significant variance components at level-2 and level-3.

In order to predict the mathematics gain scores for student i , taught by teacher j in school k , the level-1 conditional model can be expressed as follows.

$$(\text{MATHGAIN})_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{PRESCORES}) + \pi_{2jk}(\text{SES}) + e_{ijk} \quad (4)$$

where π_{0jk} is the mean gain scores for teacher j in school k , π_{1jk} and π_{2jk} are effects of pretest scores and SES respectively at student level, and the term e_{ijk} is the random effect for student i nested within teacher j and school k that is distributed normally with mean 0 and variance σ^2 . Level-2 conditional model for teachers within school can be expressed as below.

$$\begin{aligned} \pi_{0jk} &= \beta_{00k} + \beta_{01k}(\text{CERTICON}) + \beta_{02k}(\text{TCHREXP}) + r_{0jk} \\ \pi_{1jk} &= \beta_{10k} + \beta_{11k}(\text{CERTICON}) \\ \pi_{2jk} &= \beta_{20k} + \beta_{21k}(\text{CERTICON}) \end{aligned} \quad (5)$$

where β_{00k} , β_{10k} , and β_{20k} are the intercepts associated with level-2 model. Further, β_{01k} , β_{02k} , β_{11k} , β_{12k} , and β_{21k} are the slopes associated with level-2 model, and the term r_{0jk} is the random effect for teacher j nested in school k .

Level-3 model for schools can be given by Equation (6) as follows.

$$\begin{aligned}
\beta_{00k} &= \gamma_{000} + \gamma_{001} (\text{SCHLPOVERTY}) + u_{00k} \\
\beta_{01k} &= \gamma_{010} \\
\beta_{02k} &= \gamma_{020} \\
\beta_{10k} &= \gamma_{100} + \gamma_{101} (\text{ADVMTHDEG}) \\
\beta_{11k} &= \gamma_{110} \\
\beta_{20k} &= \gamma_{200} + \gamma_{201} (\text{MEANEXP}) \\
\beta_{21k} &= \gamma_{210}
\end{aligned} \tag{6}$$

After substituting the equation (6) in (5) and equation (5) in (4), the single-equation can be expressed as follows.

$$\begin{aligned}
(\text{MATHGAIN})_{ijk} &= \gamma_{000} + \gamma_{001} (\text{SCHLPOVERTY}) + \gamma_{010} (\text{CERTICON}) + \gamma_{020} (\text{TCHREXP}) + \\
&\quad \gamma_{100} (\text{PRESCORES}) + \gamma_{110} (\text{PRESCORES} * \text{CERTICON}) + \gamma_{200} (\text{SES}) + \\
&\quad \gamma_{101} (\text{PRESCORES} * \text{ADVMTHDEG}) + \gamma_{201} (\text{SES} * \text{MEANEXP}) + \\
&\quad \gamma_{210} (\text{SES} * \text{CERTICON}) + e_{ijk} + r_{0jk} + u_{00k}
\end{aligned} \tag{7}$$

Equation (7) consists of fixed portions (containing γ terms) and random portions (containing e , r , and u terms) of effects. In equation (7), the term γ_{000} represents the grand mean or mean gain scores for all schools, γ_{100} , γ_{200} , γ_{010} , γ_{020} , γ_{001} are the effects of pretest scores, SES, content-area certification, teacher experience, and school poverty, respectively. The factor γ_{110} is the interaction effect between teacher's mathematics content-area certification and students' pretest scores, γ_{101} is the interaction effect between pretest scores and advanced mathematics or mathematics education degree, γ_{201} is the interaction effect between student SES and school teacher mean experience, and γ_{210} is the interaction effect between teacher's mathematics content-area certification and SES. Further, e_{ijk} , r_{0jk} , and u_{00k} are random error terms at student, teacher, and school levels, respectively.

We have analyzed both of the above models, namely, unconditional model and conditional VAM using SAS PROC MIXED procedure (see Singer, 1998). The hypothesis was tested using

the p-values associated with the effect of individual predictors and cross-level interaction effects. Research question 2 is addressed by computing the d-type effect size for teacher and school models using the formula as follows provided by Rowan et al. (2002).

$$d = \sqrt{(\text{Variance in gain scores lying among classroom})} / \sqrt{(\text{Total student + teacher} \\ + \text{school variance in student gain scores})} \quad (8)$$

The effect size for school level model is calculated after substituting the numerator by “Variance in gain scores lying among schools” in above Equation (8).

Since the students are not placed within teachers’ classrooms randomly, and predictors at student, teacher, and school level models separately provide better estimates of variance and predictors’ effects, our best choice of statistical design to estimate the variance and predictors’ effects involves selecting the HLM technique. According to many researchers, HLM can be used as an appropriate data analysis method in such situation (Berkey, Hoaglin, Mosteller, and Colditz, 1995; Goldstein, 1995; Morris and Normand, 1992; Raudenbush and Bryk, 2002; Subedi, 2005).

Results

Table 1 provides the significant effects of individual predictors at student, teacher, and school levels and effects of their cross-level interactions using conditional VAM for predicting students’ mathematics gain scores in order to answer the research question 1. The study found the significant effects of students’ pre-test scores ($p < .0001$) and SES or socioeconomic status ($p < .0001$) on mathematics gain scores. Further, the test of hypothesis revealed that the teachers’ certification in mathematics content-area ($p = .001$) and their experience ($p = .023$) significantly predicted mathematics gain scores. A slope estimate of approximately 2.0 was found for content-area certificate that can be interpreted as the mean gain score of those middle school students

taught by teachers who hold mathematics 5-9 or mathematics 6-12 content certificates was two times higher than of those students taught by teachers who did not hold such certificates.

Likewise, school poverty ($p < .0001$) significantly impacted student mathematics gain scores, and the effect was negative. The interaction effects of teacher’s mathematics content-area certification with students’ pretest scores ($p = .001$) and SES ($p = .003$) significantly predicted mathematics gain scores at a .05 level. Further, whether or not a teacher had earned an advanced degree in mathematics or in mathematics education did not significantly impact students’ gain scores.

Table 1.

Estimated Effects of Predictors for Predicting Mathematics Gain Scores in Conditional Model

Standard Effect	Estimate	Error	t-Value	p-Value
Intercept	19.331	0.698	27.72	<.0001
Pretest Scores	0.026	0.001	37.03	<.0001
SES	-2.148	0.325	-6.62	<.0001
Content Certificate	1.973	0.615	3.21	0.001
Experience	0.042	0.019	2.28	0.023
School poverty	-4.1461	0.829	-5.00	<.0001
Pretest Scores* Content Certificate	0.003	0.001	3.30	0.001
SES* Content Certificate	0.910	0.305	2.99	0.003
Pretest Scores* Pct. Adv. Math. Deg.	0.015	0.005	3.12	0.002
SES * School Mean Experience	-0.120	0.041	-2.95	0.003

Note: “Pct. Adv. Math. Deg.” denotes percent of advanced mathematics degree.

However, interaction effect between the percent of teachers in a school with advanced degrees in such field and students’ pretest scores ($p = .002$) is found significant. Similarly, the interaction effect between the school mean teacher experience (i.e., average years of teacher experience for a

particular school) and SES ($p = .003$) impacted significantly student mathematics gain scores, and it had negative impact.

Table 2 provides the estimation of variance explained, p-value, and d-type effect sizes for predicting mathematics gain scores in unconditional and conditional VAM at teacher level. The test of hypothesis regarding “no significant teacher-to-teacher variance in mean NCE gain scores” to predict students’ mathematics gain scores, pertaining to research question 2, is rejected for both unconditional model and conditional VAM (with p-values $< .0001$). For unconditional and conditional VAM at teacher level, the d-type effect sizes were .19 (3.6% variance explained) and .22 (4.6% variance explained), respectively, with an increase of .03 in effect size for the conditional VAM.

Table 2.
Estimation of Variance Explained, Significance, and Effect Size at Teacher and School Levels

Random Effect	Variance Component	Variance Explained	p-value	Effect Size (d-type)
Teacher level effect				
Unconditional model	4.500	3.6%	$< .0001$	0.19
Conditional VAM	4.645	4.6%	$< .0001$	0.22
School level effect				
Unconditional model	0.467	0.4%	.0410	0.06
Conditional model	0.258	0.3%	.1636	0.05

Table 2 also depicts the estimation of variance explained, p-value, and d-type effect sizes for unconditional and conditional models at school level. According to the results, only the unconditional model showed significant school-to-school variance (with $p = .0410$). The d-type effect sizes, at school level, were .06 (0.4% variance explained) and .05 (0.3% variance explained)

respectively for unconditional and conditional models, with a decrease of .01 in effect size for the conditional model. Thus, a negligible effect size was found at school level for both unconditional and conditional models. Given the above results for conditional model at school level that the variance explained was not significant and a negligible amount of effect size was found, it shows that the school level factors were not important for measuring teacher effectiveness.

Discussion

The findings of this study have several implications. Discussing about student level predictors in the model, the findings showed positive impact of students' prior status scores and negative impact of their socioeconomic status on students' gain scores. It is not surprising that high achieving students in prior year will tend to have high gain scores in current year. However, since the purpose of this study is to predict student gain scores using value-added model that considers the adjustment of student level covariates (such as prior scores and background variables) at student level model, we have examined the effects of these predictors in the model. The magnitude of effect sizes at teacher level (models) using these predictors are similar to the effect sizes reported by Rowan et al. (2002).

Several teacher level factors were important for determining teacher effectiveness. For example, teacher content-area certification had significant positive impact on students' gain scores. This implies that the schools should focus on hiring the teachers who have content-area certification in mathematics in order to increase students' gain scores. This finding also concurs with the results from past research (see Darling-Hammond et al., 2005). Further, significant positive effects on gain scores were found due to interactions of content-area certification with students' pretest scores and SES. Thus, we can claim that the teachers holding content-area

certification will be able to increase student achievement gains after the interaction with student level predictors. Another piece of evidence that the teachers with content-area certification have key role in increasing students' gain scores is that this factor produced significant positive interaction effect after interacting with SES. Although SES showed significant negative effect originally (in level-1 model), it has been changed from negative to positive (effect) after the interaction with content-area certification, which is an important implication. Further, it can be claimed that more senior teachers are instrumental in increasing students' gain scores according to the findings of this study.

At school level, school poverty showed significant negative effect on students' gain scores. However, significant positive effect is found due to the interaction of percent of advanced degree in mathematics or mathematics education with pretest scores. This implies that given students' prior achievement, greater the percent of advanced degree teachers in mathematics related field in a school, the higher would be students' gain scores. School teacher mean experience showed a significant negative effect on student achievement gain while interacted with SES. This means that given the SES of students, the schools with concentrations of teachers with rich experience did not help increase students' gain scores.

Both unconditional model and conditional VAM at teacher level showed significant variances and moderate effect sizes. However, comparing the effect sizes of both models (in Table 2), the conditional VAM with predictors at teacher level model is preferable to the unconditional model since the former model produced larger effect size than the later model. At school level, both unconditional and conditional models explained small percents of variance and, consequently, the effect sizes of negligible magnitude were produced. To our surprise, the conditional VAM also showed a negligible effect size at school level even after including the factors such as school poverty and other significant interaction effects at this level. Thus, the negligible effect sizes

produced at school level indicated that school level factors do not have significant contribution for measuring teacher effectiveness.

The important finding of this research is that if we assign the teachers with mathematics content-area certification to teach impoverished students in a school, then these teachers can increase students' gain scores. However, teachers in a school with high concentration of many years of teaching experience, teaching impoverished students, did not produce effective results which could increase students' gain scores.

Conclusions

This study employed a three-level HLM using unconditional model and conditional VAM to predict mathematics gain scores in middle schools. Such models were employed in order to measure the effects of student, teacher and school level predictors and examine the magnitude of d-type effect sizes at teacher and school levels for both unconditional model and conditional VAM.

The findings indicated significant positive effects of teachers' mathematics content-area certification, teacher experience and the interaction effects of content-area certification with students' pretest scores and SES. The findings of this study imply that the teacher quality, represented by teacher content-area certification in mathematics and teacher experience as well as interaction effects associated with these predictors, are important factors in predicting mathematics gain scores in middle schools. We found that the conditional VAM produced larger effect size than that of unconditional model at teacher level. Since we found that the conditional VAM (with predictors in the model) produced larger effect size than that of unconditional model at the teacher level, the former model is preferable to use than the later model. Further, the effect sizes associated with school level model were negligible although school poverty and some other

interaction effects showed significant impact on students' mathematics gains scores. This study provided the evidence that school level factors are not important for measuring teacher effectiveness.

This research provides important information on teacher and school evaluations for schools, school districts, and the Department of Education in the states. First, given the significant effects of relevant predictors to measure teacher effectiveness, the results will be very beneficial as the potential predictors can be controlled in order to increase gain scores, and reforming schools. Second, evaluators and researchers can replicate similar conditional VAM in order to measure teacher effectiveness in their context.

Future studies are suggested to cover more grades and more school districts since this research is limited only within middle schools in a large urban school district. Researchers are also recommended to describe the teacher effectiveness based on simple effects, as demonstrated by Subedi (2005), in addition to d-type effect size.

References

- Allen, M. B. (2005). *Eight questions on teacher recruitment: What does the research say?* Denver: Education Commission of the States. Retrieved July 1, 2009 from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/b9/42.pdf.
- Ballou, D., & Podgursky, M. J. (2000). Reforming teacher preparation and licensing: What was the evidence? *Teachers College Record*, 102(1), 5-27.
- Berkey, D., Hoaglin, F., Mosteller, F., & Colditz, G.A. (1995). A random effects regression model for meta-analysis, *Statistics in Medicine*, 17, 395-411.
- Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California*. San Francisco: Public Policy Institute of California.
- Croninger, R. G., Rice, J. K., Rathbun, A. and Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review*, 26(3), 312-324
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1). Retrieved August 17, 2009 from <http://epaa.asu.edu/epaa/v8n1/>
- Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, 13(42). Retrieved March 17, 2009 from <http://epaa.asu.edu/epaa/v13n42/>
- Decker, P. T., Mayer, D. P., & Glazerman, S. (2004). *The effects of Teach for America on student: Findings from a national evaluation*. Princeton: NJ, Mathematica Policy Research Inc.
- Goldhaber, D., & Brewer, D. (1998). Why should we reward degrees for teachers? *Phi Delta Kappan*, 80, 134-138.
- Goldhaber, D., & Brewer, D. (2000). Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129-145.
- Goldhaber, D., & Brewer, D. (1998). Why should we reward degrees for teachers? *Phi Delta Kappan*, 80, 134-138.
- Goldstein, H. (1995). *Multilevel statistical models* (second edition). New York: John Wiley.
- Government Accountability Office (GAO) (2006). *State face challenges measuring academic growth that education's initiatives may help address*. Report to Congressional requesters United States. Retrieved August 27, 2009 from <http://www.ccsso.org/content/pdfs/StatesFaceChallengesGAOpaper.pdf>
- Greenberg, E., Rhodes, D., Ye, X., & Stancavage, F. (2004). *Prepared to teach: Teacher preparation and student achievement in eighth-grade mathematics*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources*, 39(2), 326-354.
- Haycock, K. (1998). *Good Teaching Matters*. How well-qualified teachers can close the gap. *Thinking K-16*, 3, 1-14. Washington DC: The Education Trust.
- Howley, C. (1996). Compounding disadvantage: The effects of school and district size on student achievement in West Virginia. *Journal of Research in Rural Education*, 12, 25-32.

- Jerald, C. D., & Ingersoll R. M. (2002). *All talk, no action: Putting an end to out-of-field teaching*. Washington, DC: The Education Trust.
- Jordan, H. R., Mendro, R., & Weerasinghe, D. (1997). *Teacher effects on longitudinal student achievement: A preliminary report on research on teacher effectiveness*. Paper presented at the National Evaluation Institute, Indianapolis, IN. Kalamazoo, MI: CREATE: Western Michigan University.
- Kane, T. J., Rockoff, J. E., Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness? Evidence from New York City*. A Technical Report. Cambridge, MA: National Bureau of Economic Research.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 38-62.
- Lippman, L., Burns, S., McArthur, E. (1996). *Urban schools: The challenge of location and poverty*. A Technical Report. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13, 125-145.
- Morris, C., & Normand, S. (1992). *Hierarchical models for combining information and for meta-analysis*. *Bayesian Statistics*, 4, 321-344.
- Pituch, K. A. (1999). Describing school effects with residual terms: Modeling the interaction between school practice and student background. *Evaluation Review*, 23(2), 190-211.
- Raudenbush, S. W., & Bryk, A. (2002). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods* (second edition). Newbury Park, CA: Sage.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Chicago: Spencer Foundation.
- Rowan, B., Correnti, R., & Miller, R. (2002). What large-scale, survey research tells us about teacher effects on student achievement; Insights from the "Prospects" study of elementary schools. *Teachers College Record*, 104(8), 1525-1567.
- Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville: University of Tennessee, Value-Added Research and Assessment Center.
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. New York: Pergamon.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24, 323-355.
- Subedi, B. R. (2007). Predicting reading proficiency in multilevel models: An ANOVA-like approach of interpreting effects. *Educational Research and Evaluation*, 13(4), 327-348.
- Subedi, B. R. (2005). *A Demonstration of three-level hierarchical generalized linear model applied to educational research*. Ph.D. dissertation, Florida State University, Tallahassee, Florida.
- Subedi, B. R., Swan, B., & Hynes, M. C. (2009). *Predicting student gains with hierarchical models: A value-added approach to measure teacher effectiveness*. Paper presented in 2009 annual meeting of American Educational Research Association, April 13-17, CA: San Diego.
- Subedi, B. R., Swan, B., & Hynes, M. C. (2005). *Middle-school teacher preparation, certification and students' mathematics achievement in a large urban district*. A Technical Report. The Multi-University Reading, Mathematics and Science Initiative. Learning System Institute, Florida State University.

- Swan, B. (2006). *Middle school mathematics teacher certification, degree level, and experience, and the effects on teacher attrition and student mathematics achievement in a large urban district*. Ph.D. dissertation, University of Central Florida, Orlando, Florida. Retrieved from Dissertations & Theses University of Central Florida-FCLA database. (Publication No. AAT 3233679).
- Tobe, P. F. (2009). Value-added models of teacher effects. In L. J. Saha, & A. G., Dworkin (Eds.), *International Handbook of Research on Teachers and Teaching*, (pp. 1113–1134). NY: Springer.
- U.S. Department of Education (USDOE), Office of Postsecondary Education. (2005). *The Secretary's fourth annual report of teacher quality: A highly qualified teacher in every classroom*, Washington, DC: USDOE.
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Educational Policy Analysis Archives*, 10(12). Retrieved March 22, 2009 from <http://epaa.asu.edu/epaa/v10n12/>