MCGILL SOCIOLOGICAL REVIEW, VOLUME 5 (July 2015): 1-18

Teacher Quality and Student Achievement: A Multilevel Analysis of Teacher Credentialization and Student Test Scores in California High Schools

RYAN SEEBRUCK
University of Arizona, Tucson, AZ

Abstract: A large body of education literature reveals positive effects of teacher quality on student achievement. However, most studies are limited to elementary schools and do not use proper regression techniques such as multilevel modeling with fixed effects, which are preferable given the nested nature of education data. I re-test the validity of past research by analyzing the relationship between teacher quality and student achievement in California high schools. I employ multilevel modeling with fixed effects on a sample of 951 high schools (Level 1) nested in 370 school districts (Level 2). Teacher quality is operationalized as the percentage of credentialed teachers. Student achievement is operationalized as the percentage of students who achieve proficiency on state-issued reading and mathematics exams. I test and find support for two hypotheses that teacher credentialization has positive effects on both measures of student achievement. Given the variance in the distribution of credentialed teachers across California, this finding is relevant to policy makers concerned with educational egalitarianism.
Key Words: teacher quality, student achievement, credentialization, multilevel regression, fixed effects

Résumé: Une grande part de la recherche en éducation révèle les effets positifs de la qualité des enseignants sur le succès des étudiants. Toutefois, la plupart des études sont limitées aux écoles primaires et n'utilisent pas les techniques de régression appropriées tel que les modèles multi-niveaux à effets fixes, qui sont préférables considérant la nature nichée des données sur l'éducation. Nous testons à nouveau la validité des recherches existantes en analysant la relation entre la qualité des enseignants et le succès des étudiants dans les écoles secondaires californiennes. Nous utilisons un modèle multi-niveaux avec des modèles à effets fixes sur un échantillon de 951 écoles secondaires (niveau 1) nichés dans 370 districts scolaires (niveau 2). La qualité des enseignants est opérationnalisée comme la proportion des enseignants accrédités. Le succès des étudiants est opérationnalisé comme la proportion d'étudiants qui ont les compétences nécessaires pour les examens gouvernementaux en lecture et en mathématique. Nous testons nos deux hypothèses : l'accréditation des enseignants a un effet positif sur les deux mesures du succès des étudiants. Celles-ci sont supportées par les résultats. Considérant la variance dans la distribution des enseignants accrédités à travers la Californie, ces résultats sont pertinents pour les décideurs publics préoccupés par l'égalité en éducation
Mots-clés: Qualité des enseignants, succès scolaire, accréditation, régressions multi-niveaux, modèles à effets fixes

Student achievement is one of the strongest predictors of future income (Hanushek 2011; Hanushek and Woessman 2009; Hanushek and Zhang 2009; Lazear 2003; Mulligan 1999), but there is a persistent achievement gap among racial and socioeconomic lines in the U.S., with minority and poor students not performing as well as their white or wealthy peers (Barton and Coley 2009). Several factors contribute to these achievement gaps, but those with the largest impact—and also those most easily remedied by policy changes—are school factors such as adjustments to curriculum, class size, availability of technology, teacher preparation and experience (Barton and Coley 2009). Teacher

quality, in particular, is seen as *the* most important variable affecting student achievement, even more so than demographic factors (Darling-Hammond 2000).

However, there is a significant discrepancy in access to high quality teachers: minority and low-income students in the U.S are less likely to be taught by certified teachers (Clotfelter, Ladd, and Vigdor 2006, 2007a, 2007b, 2010) and are more likely to be taught by inexperienced teachers (Barton and Coley 2009). This discrepancy in access to high quality teachers contributes to the achievement gap between racial and socioeconomic groups (Clotfelter et al 2010). More disconcerting is that, not only have these discrepancies in access to equal educational resources remained consistent over time, but in many cases—such as with access to certified teachers—the gaps are widening (Barton and Coley 2009). These achievement gaps persist despite an elevated priority among politicians in redressing them (Barton and Coley 2009).

To better understand the correlated discrepancies in student achievement and access to quality teachers, it is useful to re-examine the relationship between those two factors. Although there is strong theoretical support for the connection between teacher quality and student performance (see Darling-Hammond 2000; Prince 2002), Stinchcombe's (1987) assertion of the need for multiple tests of theories in order to improve their robustness and applicability holds in this case. It is beneficial to test these relationships in different contexts using different methods, especially since many extant studies in the sociology of education are focused on elementary schools and rely on correlation analyses or simple regression techniques to analyze the relationship between these variables. Given the nested nature of most social phenomena, particularly those in the education realm, a more reliable estimator of the relationship between teacher quality and student achievement requires multilevel modeling techniques—particularly those that employ fixed effects, which is necessary to mitigate "one of the most serious statistical problems associated with the measurement of teacher effectiveness, namely the fact that teachers are not randomly distributed" across schools or classrooms, and thus, are not randomly distributed across students (Clotfelter et al 2010:657).

I therefore employ multilevel regression models with fixed effects to assess the relationship between teacher quality and student performance in California high schools, considering both school-level (Level 1) and district-level (Level 2) factors. I explore two research questions. First, do schools with higher percentages of credentialed teachers have higher percentages of students who achieve proficiency on state-issued reading exams? Second, do schools with higher percentages of credentialed teachers have higher percentages of students who achieve proficiency on state-issued mathematics exams? I find support that teacher credentialization is positively associated with student performance in both reading and mathematics. This reaffirms the need for policy makers seeking to close the student achievement gaps to focus on providing all students equal access to credentialed teachers.

Theoretical Context

Teacher Quality and Student Achievement

Sociology of education researchers have long debated the impact of school resources on student performance. Some contend that schools have little independent impact compared to social context or student demographics (Coleman et al. 1966); others argue that class size (Mosteller 1995), school size (Haller and Monk 1993), and teacher quality (Ferguson 1991) all play a role in student learning (see Darling-Hammond 2000). While a plethora of background factors such as race, socioeconomic status, and home life (parental involvement, nutrition, excessive television watching, etc.) are related to student achievement (Barton and Coley 2009), research shows that school-level factors have considerable importance as well. Thus, there is a consensus among most recent education research that school-level variables do impact student performance, with particularly strong effects for teacher

characteristics (Ferguson 1991; Sanders and Rivers 1996; Jordan et al. 1997; Darling-Hammond 1997, 1999, 2000).

In fact, several studies argue that teacher quality is so important that it can explain away the achievement gap that disfavors minority and low-income students. For example, in their study of the ability of high quality teachers to close achievement gaps, Rowan et al. (2002) examined the effects of teacher quality on students of different races, gender, and socioeconomic statuses. They found that the size of achievement gaps between students of different backgrounds within the same school varied by classroom, suggesting that some teachers are more effective at closing achievement gaps between students of different backgrounds. Similarly, in a large-scale study on teacher quality and educational equality in Nevada, Borman and Kimball (2005) used multilevel models (students nested in classrooms) to show that classes taught by higher quality teachers produced higher mean achievement than those taught by lower quality teachers. Teacher quality was operationalized as experience and evaluation ratings. Their sample included nearly 5,000 elementary students and controlled for race, socioeconomic status, and student pre- and post-test scores.

In their longitudinal study of a cohort of elementary students, Sanders and Rivers (1996) found that the effects of teacher quality on student achievement are additive and cumulative. In examining student achievement data for a single cohort that spanned from second grade to fifth grade they found that student achievement at each grade was correlated positively with teacher quality. Of most interest, though, was the discovery of residual and cumulative effects: while individual students did not recover after a year under an ineffective teacher, students who spent a year under an effective teacher experienced benefits up to two years later. The authors conclude that teaching quality is more highly correlated with student achievement than other variables such as students' socioeconomic status or the racial composition of the school.

Jordan et al. (1997) obtained similar results in their longitudinal analysis of teacher quality on student performance in Dallas, Texas. Jordan found that fourth-grade students taught by highly effective teachers for three consecutive years scored thirty-five percentile points higher in reading and fifty percentile points higher in mathematics than a similar group who had been taught by a series of weak teachers. In fact, Hanushek (1992) argues that the negative consequences of even one year of instruction from an ineffective teacher can be nearly impossible to fully overcome. This leads Prince (2002:13) to conclude that "teacher quality is the single most important school variable affecting student achievement."

In sum, many sociology of education researches argue that teacher quality is one of the predominant predictors of student achievement, even more so than student background characteristics such as language background, race, ethnicity, or poverty. Certainly student background characteristics impact student achievement, and attention should be paid to addressing inequalities in those realms, but these are factors external to public education policy. In order to promote educational egalitarianism, school districts must strive to provide equal access to education resources. Such resources include expenditures per student and similar class sizes, but by far the most important resource is teacher quality.

Measuring Teacher Quality

While the literature agrees that teacher quality is important to student achievement, there is no standard definition of teacher quality, motivating some researchers to determine the relationship between various measures of teacher quality and student achievement. For instance, in a study of New York City schools, it was found that differences in teacher qualifications (educational degrees, certification status, and experience) accounted for approximately ninety percent of the total variation

in average school-level student achievement in reading and mathematics at all grade levels tested, holding student characteristics constant (Armour-Thomas et al. 1989). While it is clear that these differences in teacher qualifications matter to student achievement, it is less clear which has the most impact—something that is important to identify for school districts with constrained budgets trying to economize resources while maximizing student achievement.

While many teacher quality measures are thought to contribute to the distinction between a high and low quality teacher, certification status is often considered the most reliable and robust predictor of student achievement (Darling-Hammond 2000). To become certified as a teacher in most states requires formal education in a state-approved education program, the completion of either a major or minor in the subject field, plus minimum satisfaction of education credits and student teaching credits (Darling-Hammond 2000). Because of the strictness and thoroughness of these requirements, certification status is one of the strongest indicators of quality teacher performance.

Some in the education policy debate believe that teacher credentialization is a poor predictor of student achievement and should be discarded in favor of systems based on cognitive ability and classroom efficacy (see Walsh 2001). However, a plethora of research finds that teacher credentialization, due to the positive relationship between educational training and teacher performance, significantly affects student achievement at both the elementary and high school levels (Clotfelter et al 2006, 2007a, 2007b, 2010; Goldhaber and Anthony 2007; Rockoff 2004). These studies show that certified teachers are more effective than uncertified teachers because the latter often have difficulties with classroom management, curriculum design, teaching effectiveness, and diagnosing students' learning needs (Bents and Bents 1990; Darling-Hammond 1992). Furthermore, uncertified teachers tend to be rated lower on their teaching skills by principals, supervisors, and colleagues (Feiman-Nemser and Parker 1990). Goldhaber and Brewer (2000) found that certified teachers have a statistically significant positive impact on student test scores relative to teachers who are not certified in their subject area. Fuller (1999) found that in Texas students in districts with higher proportions of licensed teachers were significantly more likely to pass state achievement tests even after controlling for student socioeconomic status, school wealth, and teacher experience. Fetler (1999) analyzed standardized test scores of 1.3 million high school students in 785 California high schools and found that, after controlling for socioeconomic status, higher test scores correlated negatively with the number of non-credentialed teachers at a school and positively with years of teaching experience.

Clotfelter et al (2010) examined the relationship between teacher credentialization and test scores for first- and second-year high school students in North Carolina. North Carolina employs statewide end-of-course (EOC) tests at the high school level, which count for one-fourth of a student's grade in a course. Clotfelter et al examined the relationship between credentialization and EOC test scores in five subjects: algebra, economics, legal and political systems, English, geometry, and biology. They found that, overall, students taught by weak teachers were expected to perform nearly one-fourth a standard deviation lower on EOC tests than those taught by strong teachers. Considering that education is a cumulative process and that the coefficient estimates do not account for measurement error in student exams (Boyd et al 2008), the true effects of teacher credentials are likely much larger. Thus, Clotfelter et al (2010) conclude that the effect of teacher credentialization on student achievement is large and, therefore, credentialization is important enough that it should be considered in public policy decisions aimed at improving student achievement. Given the maldistribution of high quality teachers in North Carolina that disadvantage poor and minority students (Clotfelter et al 2010), public policy makers seeking to create egalitarian educations systems should take note.

In short, of the various measures of teacher quality, certification status is typically considered one of the strongest distinguishing factors affecting student achievement. However, extant literature examining the relationship between certification and student achievement has largely focused on

elementary schools or has used simple regression techniques to test the relationship. I seek to strengthen the robustness of this theorized relationship by employing multilevel regression models, which are a better fit for the nested data common to education studies.

Data, Hypotheses, and Methods

Data and Hypotheses

To examine the relationship between teacher quality and student achievement, I collected data made publicly available by the California Department of Education.[1] To maintain uniformity, I restrict my sample to high schools during the 2007-2008 school year, for a total sample size of 951 high schools in 370 school districts.[2] I analyze the effects of teacher credentialization on student achievement over two models, controlling for other independent variables cited by the literature as important predictors of student achievement. Seven of the independent variables are at the school level (Level 1) and one is at the district level (Level 2). Goodness-of-fit tests (described later) were used to eliminate non-relevant variables in order to build parsimonious models.

The outcome of 'student achievement' is represented by two dependent variables: 'reading proficiency' and 'math proficiency'. (See the Appendix for a summary of variable construction, summary statistics, and between- and within- differences in variation). Reading proficiency is operationalized as the percentage of students enrolled in a particular school for the full academic year scoring at or above the proficient level on the state-issued reading assessment tests. Math proficiency is operationalized as the percentage of students enrolled in a particular school for the full academic year scoring at or above the proficient level on the state-issued math assessment tests. The California State Board of Education administers the exams and sets its own standard for proficiency.

The main predictor of student achievement, teacher quality, is measured by 'certification status'. Certification status is a school-level (Level 1) variable measured as the percentage of teachers at a particular school who are certified. The terms certified and credentialed are used interchangeably in this report. In the state of California becoming a certified teacher involves many components: holding a bachelor's degree and completing a professional teacher preparation program (including student teaching experience) from an accredited university; passing a basic skills test in reading, writing, and mathematics issued by the state board of education; and achieving proficiency on subject-matter examination (CTC 2014). The rigor of such certification programs makes this one of the best predictors of teacher quality, and therefore the distribution of certified teachers across California schools should effect the distribution of student achievement as well. The mean percentage of certified teachers at a school is 89.3 percent (see Table B). However, with a range from 9 to 100 percent, and a standard deviation of 11.1 percent, there is considerable variance in this distribution of certified teachers across California's school, meaning that many students do not have the same opportunities to be taught by quality teachers.

I also employ seven control variables. Level 1 control variables include the total enrollment of students at the school; the average class size at the school; the percentage of students at the school who are deemed 'English learners'; the percentage of students at the school who receive free or reduced school lunch from the federal government, which is used as a proxy for socioeconomic status of a school's students; the percentage of students at the school who are non-white; and whether the

[1] California education data are available via Data Quest (http://dq.cde.ca.gov/dataquest/) and Ed-Data (http://www.ed-data.k12.ca.us/welcome.asp)

[2] The original sample size consisted of 1205 high schools, but listwise deletion of missing data—necessary for certain model fitness tests—resulted in 122 cases being dropped.

school is a charter school. The Level 2 control variable is the district's average expenditures per student. Finally, an interaction term was included for the relationship between the percentages of minority students and non-native English speakers at each school.

Based on past empirical research, I test two hypotheses:

- H_1: The percentage of credentialed teachers at a school will have a positive impact on that school's reading proficiency percentage (Model1).
- H_2: The percentage of credentialed teachers at a school will have a positive impact on that school's math proficiency percentage (Model2).

Support for these hypotheses will reaffirm the relationship between teacher quality and student achievement found by extant studies, making previous research findings more robust by increasing the applicability of the teacher quality-student achievement hypothesis to the high school level and by testing the relationship using more appropriate statistical methods (multilevel modeling using fixed effects).

Methods

Many studies examine the relationship between teacher quality and student achievement. However, most of them employ simple correlational or ordinary least squares regression techniques despite the fact that education data are multilevel. In order to decide whether a multilevel model design is necessary, one can rely on statistical, empirical, or theoretical justification (Luke 2004:17). Given the nested nature of most education data, there is strong theoretical justification for the use of multilevel modeling techniques to analyze these data. There are also several statistical techniques used to justify the use of a multilevel model. First, I run a likelihood ratio test of the multilevel models against an Ordinary Least Squares regression model. The Chi-square is highly significant for both multilevel models, suggesting the necessity of a multilevel model.[3] As further justification for using a multilevel model, I also partition the variance. Here the intraclass correlation, which can be interpreted as the proportion of total variance that is due to variation between groups (Rabe-Hesketh and Skrondal 2008), for both models indicates further justification for a multilevel model since a significant portion of variance occurs at the district level (Level 2).[4]

The risk of using single-level regression techniques on multilevel data can lead to serious errors in calculation and inference since correlations between characteristics at Level 1 do not necessarily mean the same thing as correlations at Level 2. Thus, using simple regression procedures to analyze and interpret multilevel data can lead to serious problems (Schyns 2002; Luke 2004). Two common fallacies stem out of incorrectly interpreting multilevel data at the mono-level: the ecological fallacy, which is the interpretation of group-level data at the individual level, and the atomistic fallacy, which is the interpretation of individual-level data at the group level (Schyns 2002:6). Disaggregating group-level information in a multiple regression model to the individual level lumps all of that contextual information into the individual error term, which produces correlated errors among individuals and thereby violates one of the basic assumptions of multiple regression (Luke 2004:6). Multilevel models have the ability to differentiate between effects and variation at the individual and group levels, as well as interactions between those two levels (Schyns 2002:11). In short, multilevel models produce more reliable estimates of regression coefficients because they assign variability to the appropriate level (Schyns 2002:11).

[3] Model1 had a chi2 of 88.43; Model2 had a chi2 of 73.33; both were significant beyond the 0.0000 level.

[4] The ICC for the null model of Model1 is 0.362 and for the null model of Model2 is 0.375.

Model Building

Multilevel models allow for a random intercept (the level-two residual) as well as random slopes (for level-one parameters). A random intercept model (sometimes called a variance components model) means that the starting points for the outcome variable are allowed to vary across districts (Level 2). A random slope model (sometimes called a random coefficient model) treats a level-one parameter as a random effect, meaning that its effect on the outcome variable varies between districts (Level 2) (Albright 2007). Thus, to accurately specify the models it is necessary to consider whether the covariates have fixed effects, random effects, or both (Albright 2007). Fixed effects refer to the overall expected effect of a predictor variable on the outcome variable. The random effect refers to whether this effect differs between level-two groups. A model that contains both fixed and random effects is a mixed model.

Determining which variables have random effects and which have fixed effects is accomplished by testing model fit. One method of testing model fit is transforming the measure of deviance by multiplying the natural log of the likelihood by negative two (-2LL), and then using a Chi-square goodness-of-fit test to compare nested models. However, as Luke (2004:34) points out, "One disadvantage of the deviance (-2LL) is that a model fit to the same data with more parameters will always have a smaller deviance," which conflicts with the desire for a parsimonious model. And because "a lower deviance always implies a better fit," this unfairly disadvantages nested models since they have fewer parameters (34). Another option is the Likelihood Ratio Test, but this is also limited to nested models. To compensate for this I use the Bayesian Information Criterion (BIC), a statistical criterion for model selection that imposes penalties on models with more parameters, and which can be used to compare non-nested models. To construct both models I relied on the BIC statistic to find the best combination of covariates as well as the proper specification of fixed versus random effects for those covariates.[5] For Model 1, the model with the best fit is a random intercepts model with fixed effects for eight parameters and one interaction term. For Model 2, the model with the best fit is a random intercepts model with fixed effects for seven parameters and one interaction term.

In addition to using model fitness tests to choose the appropriate type of effects in multilevel models, there is also theoretical justification for using fixed effects when analyzing certain education data. And, indeed, the results of any statistical test should be theoretically justifiable as well. Clotfelter et al (2010) point out that, when analyzing education data such as the effects of teacher credentialization on student achievement, there is potential for bias because teachers and students are not randomly assigned to classrooms. For instance, if administrators attempt to offset the potentially low scores of underachieving students by assigning them a higher proportion of higher quality teachers, the estimated effects of teacher quality would be subject to negative bias. This problem is exacerbated at the high school level because students have more freedom to choose their own courses, compared to elementary students. They address this problem by using fixed effects in their models. Because the distribution of teacher quality across California schools has also been shown, using spatial econometrics techniques, to be non-randomly distributed at the school level (Seebruck 2012), I also employ fixed effects in this paper.

[5] Diagnostics: An examination of correlation matrices and variance-inflation factors reveals no concern of collinearity. Visual analysis of residuals reveals some concern of heteroskedasticity; thus, I specify robust standard errors for both models. Outlying and influential cases were identified via Cook's Distance; the analyses were re-run with these cases removed, but since their removal did not influence the results they were ultimately included.

Results

Model 1

For the reading proficiency model, the estimated regression coefficient for the primary predictor variable, teacher credentialization, is approximately 0.25 and significant beyond the 0.000 level (see Table 1 for results). This means that every one-percent increase in teacher credentialization at a given school amounts to a 0.25-increase in the percentage of students at that given school who score above a proficient level on the state-issued reading exam, controlling for other factors such as average class size, English language ability, socioeconomic status, race, expenditure per student, school size, and charter school status. This strongly supports Hypothesis 1: *the percentage of credentialed teachers at a school will have a positive impact on that school's reading proficiency percentage*. Given the cumulative effects of teacher quality on student achievement (Boyd et al 2008; Sanders and Rivers 1996), the magnitude of the effects found here are likely to have snowballing effects throughout students' academic careers.

As expected, the percentage of English learners at a school has a highly significant negative effect on reading proficiency. Student enrollment also has a significant negative effect on reading proficiency, but the size of the coefficient is miniscule. Increased expenditures per student has a positive effect on reading proficiency, but the size of the coefficient is also miniscule. Average class size has a highly significant positive effect on reading proficiency, a counterintuitive finding that is contrary to much of the literature (Clotfelter et al 2010). In line with the literature, the main effects for both the race (percentage of minority students) and free lunch (a measure of socioeconomic status) variables have highly significant negative effects on reading proficiency. However, because there is an interaction between the percentage of minority students and the measure of socioeconomic status, the marginal fixed effects of each variable depends on the value of the other (Albright 2007). That is, the effects of race depends on socioeconomic status, and vice versa. The positive value of the interaction term means the effect of race is larger as dependence on government aid increases (or, rather, as socioeconomic status decreases).

Model 2

For the math proficiency model, the coefficient for teacher credentialization is again significant beyond the 0.000 level. The coefficient is approximately 0.32, meaning that every one-percent increase in teacher credentialization at a given school amounts to a 0.32-increase in the percentage of students at that school who score above a proficient level on the state-issued math exam, controlling for other factors. This strongly supports Hypothesis 2: *the percentage of credentialed teachers at a school will have a positive impact on that school's math proficiency percentage*. And this is in line with extant studies that also examine the relationship between teacher credentialization and student achievement at the high school level (Clotfelter et al 2010).

The other variables are all highly significant and have the same directional relationship with math proficiency that they had with reading proficiency. A notable, but not surprising, difference is that the coefficient for the percent of English learners is smaller than it was in the reading proficiency model, indicating that knowledge of English is slightly less important for math than it is for reading.

For the reading proficiency model (Model1), the estimate of the random-intercept standard deviation ($\sqrt{\psi}$) is 3.31. If we square this we get a variance of 10.96, which tells us that the intercept does vary district to district. The estimate of the level-1 residual standard deviation ($\sqrt{\theta}$) in the reading

Table 1 – Multilevel regression estimates for student achievement (two-tailed tests)
(L1 N=951; L2 N=370)

Variables	Model1 (reading) Robust coefficient	se	P-value	Model2 (math) Robust coefficient	se	P-value
	Fixed Part			Fixed Part		
credentialization %	0.246***	0.052	0.000	0.320***	0.055	0.000
average class size	0.831***	0.176	0.000	0.880***	0.164	0.000
English learners %	-0.508***	0.064	0.000	-0.383***	0.075	0.000
free lunch %	-0.523***	0.046	0.000	-0.540***	0.059	0.000
minority %	-0.203***	0.034	0.000	-0.211***	0.036	0.000
money per student	0.001*	0.000	0.013	0.001**	0.000	0.007
minority*free lunch	0.004***	0.001	0.000	0.004***	0.001	0.000
enrollment	-0.001*	0.001	0.029	---	---	---
charter school = 1	2.989	2.244	0.183	---	---	---
_constant	36.377	7.865	0.000	21.048	8.267	0.011
	Random Part			Random Part		
$\sqrt{\psi}$	Random intercept stand. dev.		3.307			5.335
$\sqrt{\theta}$	Level 1 residual stand. dev.		8.567			10.144
ϱ	Intraclass correlation coefficient		0.130			0.217
R^2	0.628 within	0.667 between	0.690 overall	0.497 within	0.526 between	0.580 overall

* p > (.05) ** p > (.01) *** p > (.001)

model is 8.567. Squaring this produces a variance of 73.39 at the school level. The intraclass correlation coefficient (ICC)—the percentage of observed variation in the dependent variable attributable to district-level (Level 2) characteristics (ϱ)—is calculated by dividing the random-intercept variance by the total variance (10.96 / [10.96+73.39] = 0.13). For the math proficiency model (Model2) we get a random-intercept variance of 28.46 and a Level 1 residual variance of 102.9, for an ICC of 0.217. That

is, there is a nonrandom distribution of student achievement across California high schools in both reading and mathematics.

Discussion

Results from two multilevel regression models indicate that teacher quality—operationalized as credentialization status—affects student achievement, with schools having higher percentages of teachers who are credentialed also having higher percentages of students achieving proficiency on both state-issued reading and mathematics examinations. Thus, in line with extant theory, the analyses in this paper reaffirm that teacher credentialization matters to student achievement and, consequently, providing equal access to certified teachers should be a goal of anyone interested in promoting educational egalitarianism.

That said, although this paper examines what is arguably one of the stronger measures of teacher quality (credentialization), there are many other variables that comprise teacher quality. Thus, policy makers should employ caution in focusing excessively on credentialization at the expense of creating balance with other measures of teacher quality because, even though teacher credentialization has strong effects on student achievement (Clotfelter et al 2006, 2007a, 2007b, 2010), it alone does not guarantee a high-quality teacher (Goldhaber and Brewer 2000). Furthermore, because the literature is unclear on the individual impacts that many other teacher quality measures have on student achievement, future research that more fully examines the magnitude of the effects of these teacher quality measures on student achievement has the potential to aid policy makers in appropriating resources more efficiently.

Nevertheless, if teacher quality is one of the predominant predictor variables affecting student performance, then presumably a more egalitarian education system would be one that more equally distributes its teachers. If that is so, then the U.S. public education system is far from egalitarian as several studies have shown that the more laissez-faire, locally-controlled education labor market in the U.S. produces a maldistribution of quality teachers that disfavors poor and minority students (see Prince 2002). In the aforementioned analysis of Nevada's Washoe County school district, Borman and Kimball (2005) found that classrooms with higher concentrations of minority, poor, and low-achieving students were more likely to be taught by teachers with lower evaluation scores. Similar within-district maldistributions have been identified in Baltimore (Lee 1998), Philadelphia (Watson 2001), Chicago (Rossi 2001; Peske and Haycock 2006), Milwaukee and Cleveland (Peske and Haycock 2006), with schools in those cities having high concentrations of poor or minority students being significantly more likely to be taught by more inexperienced, out-of-field, or less qualified teachers. As more experienced teachers, teachers teaching in their field, and credentialed teachers are associated with increased student achievement (Clotfelter et al 2010), this is a serious inequality in educational opportunities.

This maldistribution of quality teachers has persisted over decades with no signs of improvement, finally forcing Congress to realize that student achievement gaps cannot be closed without first addressing teacher distribution (Peske and Haycock 2006). With the *No Child Left Behind Act* (NCLB) of 2002, Congress demanded that states and districts address the maldistribution of quality teachers in high-poverty and high-minority schools, stipulating that for a district to continue receiving federal funds for disadvantaged students they had to develop a plan "to ensure that poor and minority children are not taught at higher rates than other children by inexperienced, unqualified, or out-of-field teachers" (ESEA Section 1111(b)(8)(c) and 1112(c)(1)(L), quoted in Peske and Haycock 2006:11). The subsequent requirement that each state submits its 'equity plan' by 2006 showed that Congress was finally serious about educational egalitarianism.

However, the NCLB offers little in the form of specific policy suggestions on how to meet its goals. Education scholars have compensated with a broad list of policy strategies: smaller classes, smaller schools, standards-based reforms, school-based reform, lengthening the school day or school year, extracurricular programs, charter schools, and even complete privatization of the public education system (see Prince 2002). However, Prince (2002:14) argues that none of the aforementioned strategies will work because the maldistribution problem is "primarily one of distribution." More money will not solve the problem either since some of the schools with the highest expenditures per student continue to have some of the lowest levels of student achievement. Prince (2002) therefore recommends that, given the overwhelming evidence that teacher quality affects student achievement, a better strategy is to redistribute teachers more equally.

In order to redistribute teachers more equally school districts need to revise current policies that are detrimental to the goals of educational egalitarianism. These maldistributions are caused by policies that encourage quality teachers to gravitate to affluent schools while less qualified teachers remain at poor schools (Prince 2002; Peske and Haycock 2006). For example, because of local funding policies, difference in community wealth contributes to salary differentials that enable schools in wealthy communities to attract and retain higher quality teachers (Prince 2002). This allows affluent districts the benefit of selecting from a larger pool of qualified applicants, which explains a good portion of inequalities between districts (Prince 2002). Throughout the U.S. there remain considerable inequities in teacher distribution within districts as well, which Prince (2002) attributes to unfavorable policies including seniority clauses that allow veteran teachers to choose their placements, policies that give the district office rather than principals the authority to hire and place teachers, and cumbersome transfer restrictions that discourage veteran teachers from transferring to low-achieving schools. Peske and Haycock (2006) also suggest scaling back seniority clauses that allow veteran teachers to choose their teaching assignments.

In short, there is a large body of research that shows minority and poor students are shortchanged when it comes to access to arguably the most important education resource: high quality teachers. Complementary research argues that the best way to close the achievement gap is to more equally distribute quality teachers. In their study on the differential effects of teacher quality on students, Goldhaber and Anthony (2007) found that teacher quality had the strongest effects on students from lower socioeconomic statuses. Accordingly, Haycock (1998) maintains that half of the achievement gap would disappear if districts simply ensured that all students had highly qualified teachers. For districts to meet the stipulations put forth by NCLB, they need to find the correct combination of policies that most equitably manages the careers of their educators while simultaneously appeasing teachers so as to prevent the best and most qualified teachers from leaving.

Ryan Seebruck *is a doctoral student in the School of Sociology at the University of Arizona in Tucson, Arizona. His research interests include social movements, technology, organizational theory, social network analysis, education, and structural inequality. He can be reached via email at seebruck@email.arizona.edu.*

Appendix

Table A – Variable Construction

Variable	Level	Construction
reading (dv)	school	[0-100] percentage of students enrolled in a particular school for the full academic year scoring at or above the proficient level on the state grade-level reading assessment tests
math (dv)	school	[0-100] percentage of students enrolled in a particular school for the full academic year scoring at or above the proficient level on the state grade-level reading assessment tests
credentialization	school	[0-100] percentage of teachers at a particular school who are certified
enrollment	school	[0-100] the total enrollment of students at the school in a given day in Oct.
ave. class size	school	[continuous] the average class size at the school
English learners	school	[0-100] percentage of students at the school deemed 'English learners'
free lunch	school	[0-100] percentage of students at the school who receive free or reduced school lunch from the federal government
minority	school	[0-100] percentage of students at the school of a race other than white
charter school	school	[binary] 1=charter school
Money per student	district	[continuous] the average expenditure per student in a given district
minority*free lunch	school	interaction term of *minority*meals*

Table B – Summary Statistics N=951

Variable	Mean	Std. Dev.	Min	Max
reading (dv)	54.44122	17.91197	5.3	98.3
math (dv)	52.83281	18.13092	4.8	97.5
credentialization	89.30625	11.14878	9.09	100
enrollment	1768.514	1016.388	32	4708
ave. class size	26.05058	3.944565	6.8	41.3
English learners	15.05447	12.28195	0.1	81.6
free lunch	42.48412	25.83632	0	100
minority	64.77298	27.0657	2.5	100
charter school	0.078864	0.269669	0	1
money per student	9159.518	1534.099	7078	22181

Table C – Distribution of Variation

Variable		Std. Dev.
reading (dv)	overall	17.91
	between	14.50
	within	11.15
math (dv)	overall	18.13
	between	14.89
	within	11.22
credentialization	overall	11.15
	between	8.093
	within	7.115
enrollment	overall	1016.39
	between	906.10
	within	655.41
ave. class size	overall	3.94
	between	3.77
	within	2.36
English learners	overall	12.28
	between	10.27
	within	8.10
free lunch	overall	25.84
	between	22.46
	within	14.18
minority	overall	27.07
	between	25.75
	within	10.89
charter school	overall	0.27
	between	0.08
	within	0.24
money per student	overall	1534.10
	between	1962.62
	within	0.00

Works Cited

Albright, Jeremy J. 2007. "Estimating Multilevel Models using SPSS, Stata, and SAS." Working Paper. May 11 2007. Accessed on 19 March 2014 <http://www.indiana.edu/~statmath/stat/all/hlm/hlm.pdf>

Armour-Thomas, Eleanor, Camille Clay, Raymond Domanico, K. Bruno, and Barbara Allen. 1989. *An Outlier Study of Elementary and Middle Schools in New York City: Final report*. NY: New York City Board of Education.

Barton, Paul E. and Richard J. Coley. 2009. "Parsing the Achievement Gap II." *Educational Testing Service*. Policy Information Report, April. Retrieved on 10 March 2014 <http://www.ets.org/Media/Research/pdf/PICPARSINGII.pdf>

Bents, Mary and Richard Bents. 1990. "Perceptions of Good Teaching Among Novice, Advanced Beginner and Expert Teachers." Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA. Cited in Darling-Hammond, 2000.

Berliner, David C. and Bruce J. Biddle. 1995. *The Manufactured Crisis: Myth, Fraud, and the Attack on America's Public Schools*. Reading, MA: Addison-Wesley.

Borman, Geoffrey D and Steven M. Kimball. 2005. "Teacher Quality and Educational Equality." *Elementary School Journal*. 106(1):3–20.

Boyd, Dan, Pam Grossman, Hamp Lankford, Susanna Loeb, and Jim Wyckoff. 2008. "Overview of Measuring Effect Sizes: The Effect of Measurement Error." Brief 2. National Center for the Analysis of Longitudinal Data in Education Research. November.

Clotfelter, Charles, Helen Ladd, and Jacob Vigdor. 2006. "Teacher Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41(4):778-820.

---. 2007a. "Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects." *Economics of Education Review* 26(6):673-82.

---. 2007b. "How and Why Do Teacher Credentials Matter for Student Achievement?" Working Paper No. 12828, NBER.

---. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45(3):655-681.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.

CTC. 2014. "Multiple Subject Teaching Credential: Requirements for Teachers Prepared in California." *State of California Commission on Teacher Credentialing*. Accessed on 9 August 2014 (http://www.ctc.ca.gov/credentials).

Darling-Hammond, Linda. 1992. "Teaching and Knowledge: Policy Issues Posed by Alternative Certification for Teachers." *Peabody Journal of Education* 67(3):123-154.

---. 1995. "Inequality and Access to Knowledge." Pp. 465-483 in *Handbook of Research on Multicultural Education*, edited by James A. Banks and Cherry A. McGee. New York: Simon & Schuster Macmillan.

---. 1997. Doing What Matters Most: Investing in Quality Teaching. New York: National Commission on Teaching & America's Future. November. Accessed on 19 March 2014 http://files.eric.ed.gov/fulltext/ED415183.pdf>

---. 1999. *Teacher Quality and Student Achievement: A Review of State Policy*

Evidence. Document R-99-1. Seattle: University of Washington, Center for the Study of Teaching and Policy. December. Accessed on 19 March 2014 <http://depts.washington.edu/ctpmail/PDFs/LDH_1999.pdf >

---. 2000. "Teacher Quality and Student Achievement:" *Education Policy Analysis* 8(1). Accessed on 19 March 2014 < http://epaa.asu.edu/ojs/article/view/392/515>

Feiman-Nemser, Sharon and Michelle B. Parker. 1990. *Making Subject Matter Part of the Conversation or Helping Beginning Teachers Learn to Teach*. East Lansing, MI: National Center for Research on Teacher Education.

Ferguson, Ronald F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal of Legislation* 28:465-498.

Fetler, Mark. 1999. "High School Staff Characteristics and Mathematics Test Results." *Education Policy Analysis Archives* 7(9):1-19.

Fuller, Edward J. 1999. "Does Teacher Certification Matter? A comparison of TAAS Performance in 1997 Between Schools with Low and High Percentages of Certified Teachers." Austin: Charles A. Dana Center, University of Texas at Austin.

Goldhaber, Dan and Emily Anthony. 2007. "Can Teacher Quality be Effectively Assessed? National Board Certification as a Signal of Effective Teaching." *Review of Economies and Statistics* 89(1):134-50.

Goldhaber, Dan and Dominic Brewer. 2000. "Does Teacher Certification Matter? High School Certification Status and Student Achievement." *Educational Evaluation and Policy Analysis* 22(2):129-45.

Haller, Emil J., David H. Monk, and Lydia T. Tien. 1993. "Small Schools and Higher-Order Thinking Skills." *Journal of Research in Rural Education* 9:66-73.

Hamacheck, Don. 1969. "Characteristics of Good Teachers with Implications for Teacher Education." *Phi Delta Kappan* 50(6):341-345.

Hanushek, Eric A. 1992. "The Trade-off between Child Quantity and Quality." *Journal of Political Economy* 100(1):84-117.

---. 2011. "The Economic Value of Higher Teacher Quality." *Economics of Education Review* 30:466-470

Hanushek, Eric A. and Ludger Woessmann. 2009. "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation." NBER Working Paper 14633. Cambridge, A: National Bureau of Economic Research.

Hanushek, Eric. A. and Lei Zhang. 2009. "Quality-Consistent Estimates of International Schooling and Skill Gradients." Journal of Human Capital 3(2):107–143.

Haycock, Kati. 1998. "Teaching and California's Future: Good Teaching Matters…A Lot. Education Trust, *Thinking K-16*, 3(2). Accessed on 19 March 2014 <http://www.cftl.org/documents/K16.pdf>

Jordan, Heather R., Robert L. Mendro, and Dash Weersinghe. 1997. "Teacher Effects on

Longitudinal Student Achievement: A Preliminary Report on Research on Teacher Effectiveness." Paper presented at the National Evaluation Institute, Indianapolis. Kalamazoo, MI: CREATE, Western Michigan University.

Lazear, Edward P. 2003. "Teacher Incentives." *Swedish Economic Policy Review* 10(3):179-214.

Lee, John. 1998. "Teacher Staffing and dDistribution Patterns for 1997 in Four Maryland LEAs." Paper presented at The Harvard Conference on Civil Rights and High Stakes K-12 Testing, December 4, 1998, New York.

Long, Scott J and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata,* 2nd Edition. Stata Press.

Luke, Douglas A. 2004. *Multilevel Modeling*. SAGE University Press: Minneapolis.

Mosteller, Frederick. 1995. "The Tennessee Study of Class Size in the Early School Grades." *The Future of Children* 5(2):113-127.

Mulligan, Casey B. 1999. "Galton versus the Human Capital Approach to Inheritance." *Journal of Political Economy* 107(S6):S184-S224.

Peske, Heather G. and Kati Haycock. 2006. "Teaching Inequality: How Poor and Minority Students are Shortchanged on Teacher Quality. Washington, D.C., Education Trust.

Prince, Cynthia D. 2002. "The Challenge of Attracting Good Teachers and Principals to Struggling Schools." American Association of School Administrators. January 2002. Accessed on 4 February 2011 <http://aasa.org/files/PDFs/Publications/challenges_teachers_principals.pdf>

Rabe-Hesketh, Sophia, and Andres Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata*. 2nd Edition. Stata Press.

Rockoff, Jonah. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review Papers and Proceedings* 94(2):247-52.

Rossi, Rosalind. 2001. "Teacher Woes Worst in Poor Schools." *Chicago Sun-Times*, October 10.

Rowan, Brian R., Richard Correnti, and Robert J. Miller. 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* 104:1525–1567.

Sanders, William L. and June C. Rivers. 1996. "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement." Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center. Accessed on 19 March 2014 <http://bulldogcia.com/Documents/Bulldog_CIA/Articles/sanders_rivers.pdf>

Scherer, Marge. 2001. "Improving the Quality of the Teaching Force: A Conversation with David C. Berliner." *Educational Leadership* 58(8):6-10.

Schyns, Peggy. 2002. "Wealth of Nations, Individual Income, and Life Satisfaction in 42 Countries: A Multilevel Approach." *Social Indicators Research* 60:5-40.

Seebruck, Ryan. 2012. "Educational Equality in the U.S.: An Exploratory Spatial Data Analysis (ESDA) of Teacher Quality Distribution in California School Districts." Presented at the International Symposium on Designing Governance for Civil Society, Keio University, February 6, Minato, Tokyo, Japan.

Stinchcombe, Arthur L. 1987. *Constructing Social Theories*. Chicago: University of Chicago Press.

Walsh. Kate. 2001. *Teacher Certification Reconsidered: Stumbling for Quality*. Baltimore: Abell Foundation.

Watson, Susan. 2001. "Recruiting and retaining teachers: Keys to improving the Philadelphia Public Schools." Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education. Accessed on 19 March 2014 http://files.eric.ed.gov/fulltext/ED480407.pdf