Available online at www.sciencedirect.com
ScienceDirect
Economics of Education Review 26 (2007) 673–682
Economics of Education Review
www.elsevier.com/locate/econedurev

Teacher credentials and student achievement: Longitudinal analysis with student fixed effects

Charles T. Clotfelter, Helen F. Ladd*, Jacob L. Vigdor

Sanford Institute of Public Policy, Duke University, Durham, NC 27708, USA

Abstract

We use a rich administrative dataset from North Carolina to explore questions related to the relationship between teacher characteristics and credentials on the one hand and student achievement on the other. Though the basic questions underlying this research are not new—and, indeed, have been explored in many papers over the years within the rubric of the "education production function"—the availability of data on all teachers and students in North Carolina over a 10-year period allows us to explore them in more detail than has been possible in previous studies. We conclude that a teacher's experience, test scores and regular licensure all have positive effects on student achievement, with larger effects for math than for reading. Taken together the various teacher credentials exhibit quite large effects on math achievement, whether compared to the effects of changes in class size or to the socio-economic characteristics of students.
Published by Elsevier Ltd.

Keywords: Education production function; Teacher quality; Teacher credentials

1. Introduction

Education researchers and policy makers agree that teachers differ in terms of quality, and that quality matters for student achievement. Despite extensive research, however, debate still rages about whether measurable teacher credentials can reliably predict either teacher quality or student achievement. We shed new light on this issue by using rich administrative data from North Carolina to explore a range of questions related to the relationship between teacher characteristics and credentials on the one hand and student achievement on the other. The teacher credentials in which we are most interested are those that can be affected in one way or another by policy.

This paper builds on our previous cross-sectional research on teacher credentials and characteristics (Clotfelter, Ladd, & Vigdor, 2006), but differs in its use of longitudinal data. These data include all North Carolina students in grades 3, 4 and 5 in years 1995–2004 for whom we can identify their teachers of math or reading. The longitudinal aspect of the data allows us to include in our models student fixed effects, which provide powerful protection against the left-out variable bias that typically plagues research of this type. Such data also permit us to explore in some detail the mechanisms through which teacher credentials exert their impacts.[1]

*Corresponding author.
E-mail address: helen.ladd@duke.edu (H.F. Ladd).

[1]Given the space constraints for this article, we are not able to describe all the details of the analysis. For those details and a broader set of results, the reader should refer to the longer version of this paper (Clotfelter, Ladd, & Vigdor, 2007).

0272-7757/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.econedurev.2007.10.002

2. Empirical framework

Although we focus here only on the findings related to teacher credentials, all the findings emerge from fully specified models of student achievement estimating using student level data. In recognition of education as a cumulative process, the standard starting point in the literature is a "value-added" model in which the learning that a student (denoted by i) brings to the classroom in year t is incorporated in the form of her achievement in the relevant subject in the previous year.[2] Specifically, we can write:

$$A_{it} = \alpha A_{it-1} + \beta_1 \mathrm{SCF}_i + \beta_2 \mathrm{SCV}_{it} + \sum_j \left[\beta_3 \mathrm{TCF}^{(j)} + \beta_4 \mathrm{TCV}_t^{(j)} + \beta_5 C_t^{(j)}\right] D_{it}^{(j)}, \quad (1)$$

where the relevant variables or vectors of variables are defined as follows:

A_{it} is achievement of student i in year t as measured by a normalized test score in reading or math.
A_{it-1} is achievement of the ith student in the prior year.[3]
TCF is a vector of teacher characteristics, such as the teacher's race and gender, that are fixed over time for any specific teacher.
TCV_t is a vector of teacher characteristics that vary over time, including, for example, years of teaching experience, attainment of higher degrees, or attainment of a particular type of license.
C_t is a vector of classroom characteristics that vary depending on the student's classroom each year. These include class size and characteristics of peers.
SCF_i is a vector of measurable student characteristics that are fixed over time, such as a student's race, gender, and age in grade 3.
SCV_{it} is a vector of student characteristics that vary over time. These include indicator variables for thing such as grade repetition or movement to a new school.
u_{it} is an error term.

In addition, $D_{it}^{(j)}$ is an indicator variable for whether the student had the jth teacher in year t. The coefficients β_1–β_5 are vectors rather than individual parameters, and α represents the extent to which knowledge persists from one year to the next.

As a model for estimating the effects of teacher credentials on student achievement (denoted by the coefficient vectors β_3 and β_4), Eq. (1) is flawed in at least two ways. First and most important, the relevant coefficients will be biased because, as we have documented in our prior research, teachers with stronger credentials tend to be matched at both the school and the classroom level with students who are more educationally advantaged (Clotfelter et al., 2006). That positive matching means that too much of the achievement of high achieving students would be attributed to the credentials of the teachers rather than to the unobserved characteristics of the students. Second, the inclusion of the lagged achievement variable on the right hand side of the equation is a problem, both because the variable is likely to be measured with error and because any correlation of achievement over time would make the variable endogenous.[4]

We address these two biases in two ways. First, we make use of our longitudinal data by replacing all the student-specific variables that do not vary over time with student fixed effects. In other words, we allow for the intercepts of the achievement equation to vary by student. Provided the effects are linear, the inclusion of these student fixed effects eliminates any bias associated with the nonrandom matching of teachers and students. This conclusion follows because their presence means that the only variation used to estimate the coefficients of interest is variation within, not across, individual students.

Second, to address the statistical problems that arise when the lagged achievement variable is included as a control variable on the right hand side of the equation, we either delete that variable or we move it to the left hand side by respecifying the dependent variable as a student's gain in achievement (i.e. as $A_{it} - A_{it-1}$). Neither solution is perfect.

[2]Hanushek (1997) and Hedges, Laine, and Greenwald (1994) provide contrasting summaries of the research literature on how teacher credentials and other inputs affect student achievement. Boardman and Murnane (1979) and Todd and Wolpin (2003) provide thoughtful analyses of model specification.

[3]For simplicity, the lagged achievement term refers here to the same subject as the dependent variable. In other analysis not reported here, we have included lagged achievement terms for both math and reading as well as squared terms for each of them, on the ground that prior achievement in both math and reading could affect current year achievement in either subject.

[4]Though the dependent variable is subject to measurement error as well, that error simply shows up in the error term and is not a source of bias.

When we delete the student's prior year achievement, under reasonable assumptions the coefficients of the teacher credentials will be biased downward. Only if knowledge did not persist at all from 1 year to the next (so that α equals zero) would there be no bias. When we move the lagged achievement variable to the left hand side of the equation, the estimated coefficients will be biased upward unless α were equal to one signifying no decay on knowledge over time.[5]

Though neither approach is perfect, we are confident that these two variations of the longitudinal model generate estimates of teacher credentials that bracket the true effects. Hence in all the results we report below, we present results from both models, which we label the levels model and the gains model, respectively. Note that the closer that α is to 1, the more valid will be the results from the gains model compared to the levels model. In results not reported here based on cross-sectional models without student fixed effects, α is estimated to be close to 0.70. Hence, we believe that the reported results for the gains model are closer to the true effects than are those from the levels model.

3. The North Carolina data

The data we use for this study are derived from administrative records maintained by the North Carolina Education Research Data Center, housed at Duke University. Student information, including their standardized test scores, are derived from student test records, and the teacher data from a state-maintained archive of personal records. Particularly relevant to this study is that North Carolina has been testing all students in reading and math from grades three to eight since the early 1990s and that the tests are closely linked to the state's Standard Course of Study. Thus, students in North Carolina are being tested on the knowledge and skills that the state wants them to know. Further, the existence of a relatively sophisticated test-based accountability system gives teachers strong incentives to teach them that material.

Crucial for this analysis is the identification of each student's actual math and reading teacher. We were able to identify with confidence math and reading teachers for at least 75% of all students in grades 3, 4 and 5 during the period 1994/1995 to 2003/2004. Far lower success rates for sixth to eighth graders prohibited us from including those grades in the analysis. In all our regressions, the dependent variable is a standardized end-of-grade test score in either reading or math for each student or the year-to-year change in that variable. Although the state reports all test scores on a development scale, we converted each scale score to a standardized score with mean of zero and standard deviation of one. This standardization makes it possible to compare test scores across grades and over time.

4. Results by teacher credential

Once again, we emphasize that the results we report here all emerge from the full model that we have just described. The levels regressions are based on about 1.8 million observations for students in grades 3, 4 and 5. The gains regressions are based on about 1 million observations and represent gains for 4th and 5th graders alone.[6] We begin with years of experience and master's degrees since they are the credentials that have most often been included in prior research.

Years of experience

We measure years of teaching experience as the number of years used by the state to determine a teacher's salary. Previous research (Clotfelter et al., 2006; Hanushek, Kain, O'Brien, & Rivkin, 2005) has shown that the returns to additional years of experience are likely to be highest in the early years of teaching. We allow for this nonlinearity by specifying years of experience as a series of indicator variables, with the base, or left-out category, being no experience. Table 1 reports the estimated effects for math and reading that emerge from the two forms of the model. Consistent with our expectations about the potential bias, in all cases the estimates from the gains model exceed those from the comparable levels model.

As expected, we find clear evidence that teachers with more experience are more effective in raising student achievement than those with less experience.

[5]For further discussion of the bias, see Clotfelter et al.(2007), Rivkin (2006) and Hanushek, Kain, and Rivkin (2006).

[6]We have also estimated levels models restricted to 4th and 5th graders. The estimated patterns are similar to those reported below for the levels model, with some of the coefficients slightly smaller. We prefer the reported results because they emerge from longer panels.

Table 1
Effects of teacher experience on student achievement, by subject and by type of model[a,b]

Base = no experience (years)	Math		Reading	
	Levels	Gains	Levels	Gains
1–2	0.057 (0.004)	0.072 (0.009)	0.032 (0.003)	0.043 (0.007)
3–5	0.072 (0.004)	0.091 (0.009)	0.046 (0.003)	0.064 (0.008)
6–12	0.079 (0.004)	0.094 (0.009)	0.053 (0.003)	0.071 (0.007)
13–20	0.082 (0.004)	0.102 (0.009)	0.062 (0.003)	0.082 (0.008)
21–27	0.092 (0.004)	0.118 (0.009)	0.067 (0.003)	0.096 (0.008)
>27	0.084 (0.005)	0.109 (0.010)	0.062 (0.004)	0.092 (0.009)

[a]The dependent variable is student achievement in math or reading normalized to a mean of 0 and a standard deviation of 1. See text for discussion of the full models of which these results are a subset. Levels refer to the model in which the dependent variable is current year achievement. Gains refer to the model in which the dependent variable is the gain in achievement during the year. All the models include student fixed effects. The entries are the estimated coefficients; the numbers in parentheses are robust standard errors.
[b]All coefficients are statistically significant at the 0.01 level.

Table 2
Teacher experience[a,b]

	Math		Reading	
	Levels	Gains	Levels	Gains
With interactions				
Teacher will stay 3+ years	−0.019** (0.007)	−0.033* (0.016)	−0.009 (0.006)	−0.022 (0.014)
1–2 years	0.059** (0.006)	0.061** (0.014)	0.029** (0.005)	0.032** (0.012)
Interact 1–2 years with stay 3+ years	0.002 (0.008)	0.023 (0.018)	0.007 (0.007)	0.020 (0.015)

*The coefficient is statistically significant at the 0.05 level.**The coefficient is statistically significant at the 0.01 level.
[a]See footnote "a" in Table 1.
[b]Based on the full models described in the text with the sample restricted to test scores pf students taught by teachers for whom we have information on whether they remained in teaching for 3 or more years. Robust standard errors are in parentheses.

For math, the benefits of experience rise monotonically to a peak of 0.092 standard deviations in the levels model and 0.119 in the gains model, with more than half the benefit occurring during the first couple of years of teaching. Similar patterns emerge for reading, but the magnitudes are somewhat smaller, a pattern that emerges for all the credentials we examine.

Though the positive results by years of teacher experience are clear and robust to various model specifications, the thorny issue remains of whether the rising returns to experience reflect improvement with experience or differentially higher attrition of the less effective teachers (Rockoff, 2004). We shed light on this issue by making use of information on teacher longevity as shown in Table 2. Specifically, we have added to the two models an indicator variable for whether or note the teacher remains a North Carolina teacher for at least 3 years and an interaction term between that variable and the indicator variable for 1–2 years of experience.[7] The negative coefficients of −0.019 and −0.033 on the indicator variable in the math equations suggests that the teachers who stay are less effective than those who leave, a pattern implying that our estimates of the returns to experience are actually underestimates. Moreover, the fact that the interaction terms, in both the math and the reading equations, are not statistically significant help rule out differential attrition as an explanation of the rising returns to experience. Hence, we conclude that the patterns of achievement gains shown in Table 1 are primarily attributable to learning from experience.

[7]We have not tried to add teacher fixed effects to either of these models with student fixed effects largely because of the technical difficulties of doing so. We have, however, included teacher fixed effects in comparable models without student (or school) fixed effects. The findings are fully consistent with those reported in the text.

Table 3
Achievement effects of graduate degrees[a]

	Math				Reading			
	Levels		Gains		Levels		Gains	
Basic model								
Graduate degree	−0.003	(0.002)	+0.002	(0.004)	−0.004**	(0.001)	−0.008**	(0.004)
By degree								
Master's degree	−0.002	(0.002)	0.003	(0.002)	−0.003*	(0.001)	−0.007*	(0.004)
Advanced degree	−0.045**	(0.012)	−0.052*	(0.025)	−0.025**	(0.010)	−0.048*	(0.020)
Ph.D.	−0.093**	(0.023)	−0.078)	(0.056)	−0.031	(0.019)	−0.021	(0.064)
Master's by time								
MA before teaching	−0.001	(0.003)	0.009	(0.007)	−0.005	(0.003)	−0.009	(0.006)
MA 1–5 years into teaching	0.004	(0.003)	0.007	(0.006)	0.001	(0.002)	−0.005	(0.005)
MA degree 5+ years into teaching	−0.010**	(0.003)	−0.008	(0.006)	−0.007*	(0.002)	−0.010*	(0.005)

*The coefficient is statistically significant at the 0.05 level.
**The coefficient is statistically significant at the 0.01 level.
See footnote "a" in Table 1.
[a]The entries in the first row are the coefficients of a single indicator variable for whether the teacher has a graduate degree. In the second two panels, the single graduate degree variable is replaced by the specified set of variables.

Graduate degrees

Our basic regressions include a single variable to indicate whether the teacher has a graduate degree of any type. Most of these degrees are master's degrees. The first row of Table 3 shows that despite the fact that teachers are rewarded for obtaining such a degree in the form of a higher salary—presumably as an incentive, at least in part, to make them a more effective teacher—having a graduate degree exerts no statistically significant effect on student achievement and in some cases the coefficient is negative. Thus, the higher pay for graduate degrees would appear to be money that is not well spent, except to the extent that the option of getting a master's degree keeps effective experienced teachers in the profession.

Further analysis is reported in the following rows. First we disaggregate the degree by type: master's, "advanced" and Ph.D. The category of "advanced" degree generally applies to graduate degrees that do not increase teacher salaries and teachers are not required to report them. Emerging from the second panel of the table is a relatively large negative effect for advanced degrees and a very small—and in half the cases not statistically significant—negative coefficient for master's degrees. The large negative coefficient on the Ph.D. variable for math is probably an anomaly given the small number of elementary school teachers with Ph.D.s.

In the third panel, we disaggregate the master's degrees by the period during which the teacher earned the degree. The estimates indicate that the teachers who received their degree prior to entering teaching or any time during the first 5 years of teaching were no less or no more effective than other teachers in raising student achievement. However, those who earned a master's degree more than 5 years after they started teaching appear to be somewhat less effective on average than those who do not have a graduate degree.

Teacher licensure

Teacher licensure also seems to matter. The state of North Carolina has many types of licenses which we have divided into three categories: regular, lateral entry and "other." Lateral entry licenses are issued to individuals who hold at least a bachelor's degree with a minimum 2.5 GPA and the equivalent of a college major in the area in which they are assigned to teach. Such teachers must affiliate with colleges and universities to complete prescribed coursework. Currently, the licenses are issued for 2 years and can be renewed for a third year. Because lateral entrants who remain in teaching eventually convert to a regular license, we include an additional variable to identify those teachers in a later year who initially entered as a lateral entrant. The "other" category includes

Table 4
National Board Certification[a,b]

	Math		Reading	
	Levels	Gains	Levels	Gains
NBCT (basic form)	0.020** (0.005)	0.028** (0.011)	0.012** (0.004)	0.012 (0.010)
Disaggregated form				
NBCT-2	0.024** (0.008)	0.055** (0.019)	0.026** (0.006)	0.038** (0.014)
NBCT-1	0.018** (0.008)	0.061** (0.017)	0.016** (0.006)	0.026* (0.013)
NBCTcurrent	0.018** (0.007)	0.046** (0.016)	0.016** (0.006)	0.035** (0.013)
NBCTpost	0.022** (0.005)	0.041** (0.010)	0.016** (0.004)	0.023** (0.008)

*The coefficient is statistically significant at the 0.05 level.

**The coefficient is statistically significant at the 0.01 level.

[a]See footnote "a" in Table 1.

[b]The entries in the first row are the coefficients of an variable indicating whether the teacher is nationally board certified. In the second panel, that variable has been replaced by the four new NBCT variables: NBCT-2 (or -1) takes on the value 1 for 2 years (or 1 year) before the teacher is certified. NBCTcurrent takes on the value 1 for a teacher the year she is certified. NBCTpost takes on the value 1 for any year after a teacher is certified. Robust standard errors are in parentheses.

a variety of provisional, temporary, and emergency licenses.

The base—or left-out category—is a teacher with a regular license. (Results not shown in tabular form.) Most clear are the negative effects on achievement for those with "other" types of provisional or emergency licenses, with the estimates ranging from −0.033 to −0.059 across the levels and gains models for math and −0.017 to −0.024 for reading. Teachers operating under a lateral entry license exhibit a statistically significant negative average effect on student achievement, but only in the levels model, and it is not clear whether that negative effect persists after the lateral entrant receives a regular license. These results for lateral entrants appear to be quite consistent with the more detailed investigation of pathways into teaching in New York State by Boyd, Grossman, Lankford, Loeb, and Wyckoff (2006). That study found that teachers with reduced coursework prior to entry into the profession often exhibited smaller initial gains than other teachers, but that the differentials were small and disappeared over time.

National board certification

North Carolina has been a leader in the national movement to have teachers board certified by the National Board for Professional Teaching Standards (NBPTS), and provides incentives in the form of a 12% boost in pay for teachers to do so. Such certification, which requires teachers put to together a portfolio and to complete a series of exercises and activities designed to test their knowledge of material for their particular field, takes well over a year and is far more difficult to obtain than state licensure.[8] As of 2004, our matched sample of math and reading teachers included about 300 board certified teachers in each of the grades 3–5.

Our basic finding is that teachers who are Board certified are more effective than those who are not (see first row of Table 4). There are two possible interpretations of this finding. One is that the National Board is identifying the most effective teachers. The other is that the process itself (or possibly the recognition associated with certification) makes the teachers more effective than they otherwise would have been. The rest of Table 4 sheds light on this issue by reporting results for four NBCT indicator variables embedded in the levels and gains basic models. The indicator variables NBCT-2 and NBCT-1 take on the value 1 two years and one year, respectively, prior to the year in which the teacher becomes certified. NBCTcurrent takes on the value 1 in the academic year in which the teacher is certified and NBCTpost represents each subsequent year that she is certified.

The positive and statistically significant coefficients for NBCT-2 indicate that the Board does indeed confer certification on the more effective

[8]For this study we obtained the names of all North Carolina certified teachers by year and grade from the NBPTS and then the NC Education Research Data Center matched those names to our information on all North Carolina teachers. Despite significant effort, we were able to match only about 90% of the teachers.

teachers, as would be appropriate to the extent that the policy goal is to reward effective teachers. These coefficients range from 0.024 to 0.055 standard deviations for math and from 0.026 to 0.038 standard deviations for reading. The fact that the NBCTcurrent and post coefficients are all lower (although not in a statistically significant sense) than the NBCT-2 coefficients provides no support for the hypothesis that the certification process makes teachers more effective than they otherwise would be. If anything, the lower coefficients on the NBCTpost variables suggest that teachers may be less effective—where effectiveness is measured by success in raising test scores—after receiving certification than before. These findings are fully consistent with those of Goldhaber and Anthony's more detailed (2005) study of National Board Certification in North Carolina for a somewhat earlier, but overlapping, time period.

Teacher test scores and quality of undergraduate institution

From the early 1960s through the mid-1990s, all elementary school teachers in North Carolina were required to take either the Elementary Education or the Early Childhood Education test. Included in the former was material on curriculum, instruction and assessment. Starting in the mid-1990s, teachers were required to take both that basic elementary test and a test that focused on content. We normalized test scores on each of these tests separately for each year the test was administered based on means and standard deviations from test scores for all teachers in our dataset, not just those in our subset of teachers matched to students.

In the basic specification, teacher test scores are simply normalized test scores averaged over all the tests taken by each elementary school teacher. As shown in the first row of Table 5, higher average test scores are associated with higher math and reading achievement, with far larger effects for math than for reading. To test for nonlinear patterns, in the second panel of the table we disaggregated the test scores into a series of indicator variables.

The results for math achievement are quite striking and exhibit some clear nonlinearity. Specifically, having a teacher at one of the extremes of the distribution has a big effect on achievement relative to having an average teacher. Referring to the results for the gains model, we see that teachers who scored 2 or more standard deviations above the average boosted student gains by 0.068 standard deviations relative to the average teacher, and teachers who scored 2 or more standard deviations below the average reduced achievement gains by 0.062 standard deviations. The overall difference between teachers at the two extremes is 0.130 standard deviations, which is far larger than the 0.060 standard deviations that would be predicted

Table 5
Teacher test scores[a]

	Math		Reading	
	Levels	Gains	Levels	Gains
Test score (basic model)	0.011** (0.001)	0.015** (0.003)	0.003** (001)	0.004* (0.002)
Non-linear average test score (S.D.s)				
⩾2	0.032** (0.010)	0.068** (0.025)	0.008* (0.007)	0.002* (0.009)
1.5 to 2	0.012** (0.005)	0.016 (0.011)	0.004 (0.004)	0.002 (0.008)
1 to 1.5	0.022** (0.003)	0.026** (0.007)	0.011** (0.002)	0.009 (0.006)
0.5 to 1	0.008** (0.002)	−0.007 (0.005))	0.001 (0.002)	−0.003 (0.004)
−0.5 to 0.5 (base)	–	–	–	–
−0.5 to −1	−0.003 (0.003)	−0.008 (0.006)	0.003 (0.002)	−0.004 (0.005)
−1 to −1.5	−0.012** (0.003)	−0.017* (0.008)	−0.005 (0.003)	−0.010 (0.007)
−1.5 to −2	−0.022** (0.005)	−0.024* (0.006)	−0.011** (0.004)	−0.016 (0.010)
⩽−2	−0.042* (0.008)	−0.062** (0.019)	−0.017* (0,007)	−0.010 (0.016)

*The coefficient is statistically significant at the 0.05 level.

**The coefficient is statistically significant at the 0.01 level.

See footnote "a" in Table 1.

[a]The entries in the first row are the coefficients of the average normalized teacher test score variable. Subsequent entries are based on the same models but with the single variable replaced by the series of indicator variables. Robust standard errors are in parentheses.

from the linear specification. A similar nonlinear pattern emerges from the levels model for math, but the difference between the extremes is far smaller at 0.074 standard deviations. For reading, all the effects are much smaller and any nonlinearities are hard to detect.

We also included a measure of the competitiveness of the teacher's undergraduate institution since that is a common measure of teacher credentials used in other studies. Following standard practice in the US literature, we assigned to each teacher's undergraduate institution a competitive ranking based on information for the 1997–1998 freshman class from the Barron's College Admissions Selector. Perhaps because of the rich set of other measures included in this study, and in particular the inclusion of the teacher's test score, these variables exhibit only very small effects, at most, on student achievement. In particular coming from an elite and very competitive institution apparently does not make a teacher any more effective on average relative to teachers from other institutions.

5. Interpreting the magnitudes

Each teacher brings to the classroom a bundle of personal characteristics and credentials. Hence, we illustrate the magnitudes of the estimated teacher effects by comparing teachers with different bundles of attributes. Consider for example a baseline teacher with the following relatively typical attributes listed in the first column of Table 6. The teacher has 10 years of experience, attended a competitive undergraduate college, and has a regular license, an average test score, and a graduate degree. In addition, we assume somewhat less typically that she is National Board certified. We then compare her to a teacher with far weaker credentials as described in the second column. That teacher has no teaching experience, attended a non-competitive undergraduate college, does not have a regular license, has a test score one standard deviation below average, does not have an advanced degree and is not Board certified.

Based on the lower bound estimates from the levels model and the upper bound estimates from the gains model, the following two columns depict the reasonable range of differential effects on student achievement in both math and reading, all other factors held constant. We remind the reader that the true effects are likely to be somewhat closer to those from the gains model than to those of the levels model. The first observation to emerge from the calculations is the far larger adverse effects from having a teacher with weak credentials on math than on reading achievement. For math, the total effects of having the weak teacher range from −0.150 to −0.206 standard deviations and for reading from −0.081 to −0.120 standard deviations. Second, the biggest differentials are associated with experience and licensure status. Of course, by assuming the subject teacher has no teaching experience we have magnified the effects of experience. If, instead, the subject teacher had 1 or 2 years of experience, the total effect in math would have ranged from −0.093 to −0.134 and for reading

Table 6
Interpreting the magnitudes[a]

Baseline teacher	Comparison teacher (weak credentials)	Difference in achievement (lower and upper bound estimates)[a]			
		Math		Reading	
		Levels	Gains	Levels	Gains
Ten years of experience	No experience	−0.079	−0.094	−0.053	−0.072
Competitive undergraduate college	Non competitive undergraduate college	−0.007	−0.010	*	*
Regular license	Other license	−0.033	−0.059	−0.017	−0.024
Licensure test score is average	Licensure test is 1 S.D. below the average	−0.011	−0.015	−0.003	−0.004
Graduate degree	No graduate degree	*	*	+0.004	+0.008
National Board Certified	Not National Board Certified	−0.020	−0.028	−0.012	−0.012
Total difference		−0.150	−0.206	−0.081	−0.120

*Coefficient is not statistically significant.
[a]Subject teacher minus baseline teacher; all the entries are reported in previous tables or in the text.

would have been −0.049 to −0.077. Though the comparison in Table 6 is merely illustrative, it does provide some information with which to evaluate the magnitude of the estimated effects. The question is whether these teacher effects are large or small.

Relative to the estimated effects of class size, the effects of teacher credentials appear to be quite large. Based on estimated coefficients for the class size variable in comparable models (results not shown here), an increase of five students in an elementary school class would reduce student achievement in math by about 0.015–0.025 standard deviations and in reading by about 0.010–0.020 standard deviations, far smaller effects than those associated with having a teacher with weak credentials.

It is also instructive to compare our estimated effects of teaching credentials to estimates of the impact of overall teacher quality that emerge from models that substitute teacher-specific fixed effects for explicit measures of credentials (Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Using such a model, Rockoff (2004) reports that a 1 standard deviation difference in the quality of a student's teacher (based on the distribution of teacher quality that emerges from the teacher fixed effects in his model) generates a 0.1 standard deviation difference in reading and math scores. We cannot directly compare our findings to his because of the difficulty of quantifying in terms of standard deviations the difference between our baseline and comparison teachers. Nonetheless, it seems safe to conclude from our results that measurable credentials account for a significant portion of the total effect of teacher quality on student achievement. To see why, consider the following example based on the relatively extreme assumption that our comparison teacher differs from the baseline teacher by 4 standard deviations in terms of the distribution of some overall latent quality measure. The Rockoff estimate would then predict that student achievement levels would differ by 0.4 standard deviations. Given that our analysis predicts achievement differences of the order of 0.1–0.2 standard deviations, we would conclude that the measurable credentials of teachers accounts for one-quarter to one-half of the overall effect of teacher quality on student achievement. A less extreme assumption, say of a 3 standard deviation difference between our baseline and comparison teachers, would imply that credentials account for one to two-thirds of the total effect of teacher quality.

An alternative comparison is to the effects of demographic characteristic such as parental education. From the results for student characteristics from an alternative specification of the model without student fixed effects but not included here, we find that, relative to having a parent who is a college graduate, having a parent without a college degree reduces predicted math achievement by about 0.11 standard deviations if the parent has a high school degree and by another 0.11 standard deviations if the parent is a high school drop-out. The effects are slightly larger for reading: 0.11 for those with high school degrees and another 0.14 for high school drop-outs. Thus, for math, having a teacher with weak credentials has negative effects generally comparable in size to those associated with having poorly educated parents. For reading, the negative effects associated with having a teacher with poor credentials, though still harmful for achievement, are not as harmful as having poorly educated parents.

Thus, we conclude that a variety of teacher credentials matter for student achievement and that the effects are particularly large for achievement in math. As a result, how teachers with differing qualifications are distributed among classrooms and schools matters. To the extent that the teachers with weaker credentials end up in classrooms with the more educationally disadvantaged children, schools would tend to widen, rather than reduce, the already large achievement gaps associated with the socioeconomic differences that students bring to the classroom. In related research, we have documented that such disparities exist both by race of the student and poverty level of the school (Clotfelter, Ladd, & Vigdor, 2002; Clotfelter et al., 2006; Clotfelter, Ladd, Vigdor, & Wheeler, 2006).

Acknowledgments

This paper is a revised version of a paper initially prepared for a conference organized by the World Bank on "The Contribution of Economics to the Challenges Faced by Education," Dijon, France, June 2006. The authors are grateful to the Spencer Foundation for funding for this project, the North Carolina Education Research Data Center for the data, and to Aaron Hedlund for excellent research assistance. Other research assistants, including Trip Stallings and Roger Aliaga-Diaz, contributed to earlier stages of this research project. We are grateful to them all.

References

Boardman, E. E., & Murnane, R. J. (1979). Using panel data to improve estimates of the determinants of educational achievement. *Sociology of Education*, *52*, 113–121.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy*, *1*(2 (Spring)), 176–216.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2002). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, *24*, 377–392.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, *XLI*(4 (Fall)), 778–820.

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Wheeler, J. (2006). *High poverty schools and the distribution of principals and teachers*. Sanford Institute working paper. 〈www.pubpol.duke.edu/research/papers/SAN06-08.pdf〉.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). *How and why teacher credentials matter for student achievement*. National Bureau of Economic Research working paper.

Goldhaber, D., & Anthony, E. (2005). *Can teacher quality be effectively assessed*? Updated version of an urban institute working paper.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, *19*(2), 141–164.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005, January). *The market for teacher quality*. Stanford University (unpublished).

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2006, March). New evidence about Brown v. Board of education: The complex effects of school racial composition on achievement (unpublished).

Hedges, L. V., Laine, R., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, *23*(April), 5–14.

Rivkin, S. (2006, January). Cumulative nature of learning and specification bias in education research (unpublished).

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica*, *79*, 418–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review Papers and Proceedings*.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *Economic Journal*, *113*(February), F3–F33.