
AMERICAN INSTITUTES FOR RESEARCH®

# Effects of National Board Certified Teachers on Student Achievement and Behavioral Outcomes: Studies Conducted in Two States

## Working Paper

David Manzeske
So Jung Park
Feng Liu
Trisha Borman
Natalya Gnedko-Berry
Benjamin West
Evelyn Deng

FEBRUARY 2017

# Effects of National Board Certified Teachers on Student Achievement and Behavioral Outcomes: Studies Conducted in Two States

## Working Paper

**February 2017**

**David Manzeske**
**So Jung Park**
**Feng Liu**
**Trisha Borman**
**Natalya Gnedko-Berry**
**Benjamin West**
**Evelyn Deng**


1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

**www.air.org**



8625_02/17

# Contents


Page

Acknowledgments ..... 1

Funding Disclosure ..... 2

Abstract ..... 3

Introduction ..... 4

- Previous Research on NBCTs ..... 4
- Current Studies ..... 6

Methods ..... 6

- Data ..... 7
- Samples ..... 7
- Propensity Score Matching ..... 8
- Checking Baseline Equivalence ..... 11
- Outcome Measures ..... 13
- Analytic Approach ..... 15

Results ..... 15

- Study 1: North Carolina ..... 16
- Study 2: Kentucky ..... 19

References ..... 23


# Tables


Page

Table 1. Variables to Be Used in Propensity Score Methods for North Carolina ..... 9

Table 2. Variables to Be Used in Propensity Score Methods for Kentucky ..... 10

Table 3. Standardized Mean Differences Between Treatment and Comparison Schools for North Carolina ..... 11

Table 4. Standardized Mean Differences Between Treatment and Comparison Schools for Kentucky. ..... 12

Table 5. Descriptive Statistics of 2014–15 North Carolina Student Outcomes, by Grade and Treatment Status ..... 14

Table 6. Descriptive Statistics of 2014–15 Kentucky Student Outcomes, by Grade and Treatment Status ..... 14

Table 7. Estimates of the Effects of Being Taught by NBCTs on North Carolina Student Academic Outcomes, by Grade ..... 16

Table 8. Estimates of the Effects of Being Taught by NBCTs on North Carolina Student Behavioral Outcomes, by Grade ..... 16
Table 9. Estimates of North Carolina Subgroup Analyses ..... 18
Table 10. Estimates of the Effects of Being Taught by NBCTs on Kentucky Student Academic Outcomes, by Grade ..... 19
Table 11. Estimates of the Effects of Being Taught by NBCTs on Kentucky Student Behavioral Outcomes, by Grade ..... 19
Table 12. Estimates of Kentucky Subgroup Analyses ..... 21

# Acknowledgments

We thank the North Carolina Education Research Data Center (NCERDC) and the Kentucky Center for Education and Workforce Statistics (KCEWS) for providing the data used in this work. We thank the National Board for Professional Teaching Standards (NBPTS) for providing support and information regarding National Board Certification. We also thank AIR database specialist Thomas Fink for assisting with data management solutions. Finally, we thank AIR researchers, Andrew Swanlund and Ryan Williams, for helpful comments. Any and all errors are the responsibility of the study authors, and the statements expressed are those of the authors and should not be attributed to American Institutes for Research, NBPTS, or the funding agency (U.S. Department of Education).

# Funding Disclosure

This work was funded by the National Board for Professional Teaching Standards through a 2015 U.S. Department of Education Supporting Effective Educator Development grant.

# Abstract

The purpose of this work was to examine the effect of classroom teachers who earn certification as a National Board Certified Teacher (NBCT) by the National Board for Professional Teaching Standards on mathematics and reading achievement and behavioral outcomes (attendance and discipline referrals) for students in Grades 4 and 5. Two studies were conducted separately to examine the effect of NBCTs in Kentucky and North Carolina, which are two states with relatively high concentrations of NBCTs. In each study, propensity score matching was used to match students of NBCTs to similar students of non-NBCTs from the same schools. Within each grade and within each study, propensity score matching yielded similar student groups (those taught by NBCTs and those not taught by NBCTs) on observed characteristics such as prior-year outcomes and demographics. For each grade level, the academic and behavioral outcomes of students of NBCTs were compared with the outcomes of the matched students taught by non-NBCTs. In North Carolina, there were no statistically significant student achievement differences at either grade between students of NBCTs and students of non-NBCTs. However, Grade 5 students of NBCTs had higher attendance rates than students of non-NBCTs by 0.02 standard deviations. There were no statistically significant effects on the likelihood of a student receiving an in-school or out-of-school suspension in North Carolina. In Kentucky, Grade 5 students taught by NBCTs scored higher than students of non-NBCTs on state student achievement in mathematics and reading by 0.06 standard deviations. There were no statistically significant differences in the behavioral outcomes between the student groups at either grade level in Kentucky.

# Introduction

It is important to identify effective teachers because effective teaching has consequences for student achievement and beyond. Studies continue to demonstrate that teacher effectiveness is a significant contributor to student academic achievement (Chetty, Friedman, & Rockoff, 2014a; Hanushek, 2010; Nye, Konstantopoulos, & Hedges, 2004; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). There is evidence that students who had more effective teachers are more likely to have higher attendance (Gershenson, 2016) and to attend college, have higher salaries, and be less likely to become pregnant as a teenager (Chetty, Friedman, & Rockoff, 2014b). This evidence reinforces the efforts of districts to identify, recruit, and retain the most effective teachers with respect to their effect on student outcomes as part of a strategy for increasing student learning and promoting other beneficial outcomes.

The identification of effective teachers remains a challenge, and efforts have taken myriad forms. Policies and practices vary widely as states and school districts try to develop procedures that will allow them to identify, recruit, and retain an effective teacher for each classroom. One such practice mirrors that used in medicine: board certification. In medicine, board certification acts as a trustworthy, universal signal of ensuring physician effectiveness for all of their patients.

Using a medical model, the National Board for Professional Teaching Standards (NBPTS) has established and maintained definitive standards of effective teaching, much like physicians' board standards. Since 1987, more than 112,000 teachers in all 50 states and the District of Columbia have achieved status as a National Board Certified Teacher (NBCT). For National Board certification, teacher pedagogy and content knowledge are evaluated against a set of standards established by the National Board such that certification criteria and status are identical across the country. To our knowledge, it is the only teacher certification system that classifies teachers in this way. Other markers of effective teachers are neither consistent in their criteria nor national in scope. For example, state teacher-of-the-year awards often are based on self-nominations and recognition criteria, which vary across states. Similarly, teacher evaluation systems vary between and within states, and this variation makes it difficult to identify which teachers are most effective.

The consistency and presumed quality of National Board certification is evidenced by the fact that in 33 states across the country, either state or local education agencies have policies to provide stipends, salary increases, or other incentives to teachers for achieving NBCT status and serving in their schools (NBPTS, 2015). For example, in North Carolina, NBCTs are placed on a salary schedule that is 12 percentage points above base pay and first-time candidates receive a $1,900 loan to pursue certification. In Kentucky, NBCTs receive a $2,000 annual stipend, and numerous districts pay for candidate fees. Policies such as these have their origin in the belief that NBCTs outperform their non-NBCT peers. Such a belief would justify a policy to elevate NBCT status and to motivate districts and schools to recruit and retain NBCTs.

## Previous Research on NBCTs

Several studies have examined the effect of NBCTs on student achievement relative to non-NBCTs. Such studies include one experimental study, conducted in Los Angeles Unified School

District, and other observational studies that used statewide data from North Carolina, Florida, and Washington.[1] Evidence suggests that, on average, NBCTs outperform non-NBCTs, as measured by student performance on state assessments, although not all studies report statistically significant findings.

The experimental study conducted in Los Angeles (in which students were randomly assigned to 99 teacher pairs) found no statistically significant effects of NBCTs on student achievement outcomes compared with teachers who had not pursued Board certification (Cantrell, Fullerton, Kane, & Staiger, 2008). However, in a nonexperimental comparison with students of unsuccessful applicants, students of successful applicants scored 0.22 and 0.19 standard deviations (SDs) higher in mathematics and reading, respectively.

Four observational studies have used North Carolina data with mixed results. In one of the earliest studies that assessed the effect of NBCTs on student achievement, Grade 3–5 data from 1996–97 through 1998–99 were used (Goldhaber & Anthony, 2007). The authors found that NBCTs were more effective than non-NBCTs in reading by 0.02 SDs, and there were no statistically significant findings for mathematics. Like the Cantrell et al. (2008) study, successful applicants were more effective than unsuccessful applicants by 0.13 and 0.07 SDs in mathematics and reading, respectively. Another North Carolina study used Grade 3–5 data from 1995–96 through 2003–04 (Clotfelter, Ladd, & Vigdor, 2007). NBCTs were found to be more effective than non-NBCTs in mathematics and reading, by 0.02 and 0.01 SDs, respectively. In a supplemental study that used Grade 5 data from 2000–01, NBCTs were found to be more effective than non-NBCTs in reading by between 0.03 and 0.45 SDs, although these effects were not replicated in more cautious models (Clotfelter, Ladd, & Vigdor, 2006). The fourth North Carolina study used Grade 10 data from 1999–2000 through 2002–03 (Clotfelter, Ladd, & Vigdor, 2010). NBCTs were found to be more effective than non-NBCTs by 0.05 SDs.

Two observational studies have used Florida data. Using elementary, middle, and high school data from 2000–01 through 2003–04, NBCTs were not found to be more effective than non-NBCTs (Harris & Sass, 2009). Using Grade 4–8 data from 2001–02 through 2008–09, NBCTs were found to be more effective than non-NBCTs in mathematics and reading that ranged from 0.02 to 0.03 SDs (Chingos & Peterson, 2011).

The most recent observational study was conducted in Washington. The study used Grade 4–8 data from 2006–07 through 2012–13, and the authors found that NBCTs were more effective than non-NBCTs (Cowan & Goldhaber, 2016). Specifically, NBCT elementary teachers were more effective in mathematics and reading, by 0.04 and 0.03 SDs, respectively. And NBCT middle school teachers were more effective in mathematics and reading by 0.05 and 0.02 SDs, respectively.

[1] According to information provided by NBPTS, North Carolina, Washington, and Florida have the largest percentages of current teachers who are NBCTs: 20%, 10%, and 9%, respectively.

## Current Studies

The current work consists of two studies in which students of NBCTs are compared with students of non-NBCTs. Each study corresponds to teachers and students from North Carolina and Kentucky.

The current studies use data from Grade 4 and Grade 5 students in states with some of the largest percentages of teachers who are NBCTs—20% in North Carolina and 3% in Kentucky. For context, 43 states have a smaller percentage of active NBCTs, according to information provided by NBPTS.

The current studies diverge from prior work in several ways. First, prior work used value-added modeling approaches in which all students (from within particular grade levels) were included in the analyses, whereas we use a propensity score matching approach to compare students of NBCTs with a restricted sample of only similar students of non-NBCTs. Second, we use data from more recent elementary school cohorts—we assess outcomes from the 2014–15 school year. Finally, we examine both achievement and behavioral outcomes. Specifically, achievement on Common Core-aligned mathematics and reading state assessments, and behavioral outcomes including attendance rate and whether a student received an in-school or out-of-school suspension.

In each study, we answer three primary and secondary research questions. The two primary research questions are as follows:

1. What is the effect of being taught by NBCTs on student mathematics achievement and reading achievement?
2. What is the effect of being taught by NBCTs on student attendance rate and whether a student received an in-school or out-of-school suspension?

The secondary, and exploratory, research question is:

3. Do effects vary based on student characteristics (e.g., race/ethnicity, English language learner [ELL] status)?

In what follows, we present our methods, results, and a discussion of the findings. The Methods section contains information that corresponds to both the North Carolina and Kentucky studies. The Results section contains study-specific findings. Following the results is a single section in which we discuss the findings.

# Methods

The subsections that follow are related to the data used, study samples, propensity score matching and checking of baseline equivalence that is conducted for both studies, outcome measures, and the analytic approach that is used for both studies.

## Data

We used data from multiple sources. The primary data sources are student and teacher administrative records from the North Carolina Education Research Data Center and the Kentucky Center for Education and Workforce Statistics from the 2013–14 and 2014–15 school years for students in Grades 3–5. Additionally, we obtained information on school and district characteristics from the Common Core of Data (CCD). These records include identifiers unique to each teacher and student, and they note the school, district, and year to which the data elements correspond. The identifiers allowed the research team to link students, teachers, schools, and districts across years.

## Samples

The study samples include elementary-level students in fourth and fifth grades from the 2014–15 school year in North Carolina and Kentucky. We opted to restrict our study to this group because elementary-level students are most likely to have only one teacher in a self-contained classroom who teaches multiple subjects per grade, thus enabling us to link students' records to their teacher. Students in the third grade were not included because valid pretest scores (obtained at the end of the second grade) were not available, and those in the sixth grade and above were not included because they were linked with several different teachers both within and across subject areas. Of these students, we excluded those with missing outcome (2014–15) or prior-year (2013–14) data to retain all students in the outcomes analyses. Then, we restricted the sample schools to those with at least one NBCT and one non-NBCT, assuming that schools having at least one NBCT might be systematically different from those without NBCTs.[2]

For North Carolina, of the 118,733 student observations (63,680 fourth graders and 55,053 fifth graders) from the 2014–15 school years, we could match 112,408 (60,321 fourth graders and 52,087 fifth graders; 95% match rate) with their 2013–14 data. Of these, we matched 108,678 (58,337 fourth graders and 50,341 fifth graders; 97% match rate) student observations within 1,404 schools that included valid scores and other key variables for both school years. Fourth graders are from 1,350 schools and fifth graders are from 1,326 schools. Of these fourth graders within 1,350 schools, 20,265 students from 556 schools with no NBCTs and 181 students from nine schools where all teachers are NBCTs were omitted. Of these fifth graders within 1,326 schools, 37,526 students from 587 schools with no NBCTs and 290 students attending 13 schools where all teachers are NBCTs were omitted. The final fourth-grade sample includes 37,891 students within 785 schools and 108 districts, and the final fifth-grade sample includes 32,971 students within 726 schools and 109 districts. Therefore, our final fourth-grade and fifth-grade samples each include 60% of the available data.

For Kentucky, of the 102,783 student observations (51,335 fourth graders and 51,448 fifth graders) from the 2014–15 school years, we could match 97,015 (48,390 fourth graders and 48,625 fifth graders; 94% match rate) with their 2013–14 data. Of these, we matched 93,962

[2] For example, in North Carolina, schools having only NBCTs have higher average academic performance, a lower percentage of students eligible for free or reduced-price lunch (FRPL), and a lower percentage of minority students than schools that have both NBCTs and non-NBCTs. In contrast, schools that have only non-NBCTs have lower average academic performance, a higher percentage of students eligible for FRPL, and a higher percentage of minority students than schools without NBCTs.

(46,821 fourth graders and 47,141 fifth graders; 97% match rate) student observation within 732 schools that included valid scores and other key variables for both school years. Fourth graders are from 714 schools and fifth graders are from 709 schools. Of these fourth graders within 714 schools, 38,031 students from 605 schools with no NBCTs and 82 students from two schools where all teachers are NBCTs were omitted. Of these fifth graders within 709 schools, 37,645 students attending 605 schools with no NBCTs and 142 students from four schools where all teachers are NBCTs were omitted. The final fourth-grade sample includes 8,708 students within 107 schools and 60 districts, and the final fifth-grade sample includes 9,354 students within 100 schools and 49 districts. Therefore, our final fourth-grade and fifth-grade samples include 17% and 18% of the available data, respectively.

## Propensity Score Matching

To examine the effects of NBCTs on students' academic and behavioral outcomes, we used a propensity score matching approach to identify comparison groups (students taught by non-NBCTs) that are similar to treatment groups (students taught by NBCTs) on observable characteristics. To balance the treatment and comparison groups in terms of observable characteristics, a logistic regression model was applied. By using a logistic regression model, we estimated every student's conditional probability of being assigned to NBCTs as a function of pretreatment personal, school, and district characteristics. The general form of the logistic regression model used for matching students is as follows:

$$\text{logit}(P(\text{NBCT}_i)) = \alpha + \sum_{k=1}^{k} \lambda_k Z_{ki} + \sum_{\rho=1}^{\rho} \theta_\rho C_{\rho i}$$

where $\text{NBCT}_i$ is an indicator of whether student $i$ was taught by an NBCT teacher; NBCT is 1 if student $i$ was taught by a NBCT teacher and 0 if otherwise; $P(\text{NBCT}_i)$ is the probability of student $i$ taught by a NBCT teacher; $\alpha$ is intercept; $Z_{ki}$ is a set of prior outcomes for student $i$ (e.g., students' prior mathematics and reading test scores, and prior year suspension information[3]); $C_{\rho i}$ is a set of student demographic characteristics (e.g., gender, socioeconomic status), teacher characteristics (e.g., degree, certification), and school academic and demographic characteristics (e.g., average academic performance, percentage of students eligible for free or reduced-price lunch) for student $i$; $\lambda$ is a set of coefficients that represents the association between each prior student outcome and the logit of the propensity score (PS)[4]; and $\theta$ is a set of coefficients that represents the association between each characteristic and the logit of the PS. All baseline covariates were measured before the 2014–15 treatment year to ensure that the measurements were not influenced by the treatment. For lists of North Carolina and Kentucky variables that were used for estimating PSs, see Tables 1 and 2, respectively.

[3] Kentucky prior-year student attendance data were not available.

[4] The logit of the PS is equal to $\log(\frac{P}{1-P})$.

**Table 1. Variables to Be Used in Propensity Score Methods for North Carolina**

| Variable Description | Source |
|---|---|
| **Treatment Indicator [2014–15]** | |
| Students taught by NBCTs during 2014–15 = 1, students taught by non-NBCTs during 2014–15 = 0) | NCERDC |
| **Baseline Covariates [2013–14]** | |
| *Student Characteristics* | |
| Reading pretest | NCERDC |
| Mathematics pretest | NCERDC |
| In-school or out-of-school suspension experience, a binary variable indicating whether a student received an in-school or out-of-school suspension during the school year | NCERDC |
| Student attendance rate, calculated as the number of days in attendance divided by the number of days enrolled | NCERDC |
| Free or reduced-priced lunch status, coded 1 if a student qualified for free or reduced-price lunch (0 if not) | NCERDC |
| Female, coded 1 if a student is female (0 if male) | NCERDC |
| Racial minority, coded 1 if a student is non-White (0 if White) | NCERDC |
| Limited English proficiency, coded 1 if a student is entitled to limited English proficiency resources (0 if not) | NCERDC |
| Student learning disability status, coded 1 if a student is in a special education program (0 if not) | NCERDC |
| Student taught by an experienced teacher, coded 1 if a student's teacher has more than five years' experience (0 if not) | NCERDC |
| Student taught by a licensed teacher, coded 1 if a student's teacher is fully certified (0 if not) | NCERDC |
| Student taught by a teacher with an advanced degree, coded 1 if a student's teacher obtained master's degree or higher (0 if not) | NCERDC |
| *School Characteristics* | |
| School previous year performance | CCD |
| School size | CCD |
| Student-teacher ratio | CCD |
| Percentage of minority (i.e., non-White) students in school | CCD |
| Percentage of student who are eligible for free or reduced-price lunch in school | CCD |
| Percentage of experienced teachers | NCERDC |
| Percentage of licensed teachers | NCERDC |
| Percentage of advanced degree teachers | NCERDC |

*Note.* CCD = Common Core of Data; NCERDC = North Carolina Education Research Data Center.

**Table 2. Variables to Be Used in Propensity Score Methods for Kentucky**

| Variable Description | Source |
|---|---|
| **Treatment Indicator [2014–15]** | |
| Students taught by NBCTs during 2014–15 = 1, students taught by non-NBCTs during 2014–15 = 0) | KCEWS |
| **Baseline Covariates [2013–14]** | |
| *Student Characteristics* | |
| Reading pretest | KCEWS |
| Mathematics pretest | KCEWS |
| In-school or out-of-school suspension experience, a binary variable indicating whether a student received an in-school or out-of-school suspension during the school year | KCEWS |
| Free or reduced-priced lunch status, coded 1 if a student qualified for free or reduced-price lunch (0 if not) | KCEWS |
| Female, coded 1 if a student is female (0 if male) | KCEWS |
| White, code 1 if a student is White (0 if non-White) | KCEWS |
| Black, code 1 if a student is Black (0 if non-Black) | KCEWS |
| Hispanic, code 1 if a student is Hispanic (0 if non-Hispanic) | KCEWS |
| Other racial group, code 1 if a student is in a racial group other than White, Black and Hispanic (0 if is one of the three racial groups) | KCEWS |
| Limited English proficiency, coded 1 if a student is entitled to limited English proficiency resources (0 if not) | KCEWS |
| Student learning disability status, coded 1 if a student is in a special education program (0 if not) | KCEWS |
| Student taught by an NBCT, coded 1 if a student's teacher is an NBCT (0 if not) | KCEWS |
| Student taught by a teacher with a bachelor degree, coded 1 if a student's teacher obtained a bachelor's degree (0 if not) | KCEWS |
| Student taught by a teacher with a master's degree, coded 1 if a student's teacher obtained a master's degree (0 if not) | KCEWS |
| Student taught by a teacher with other degrees, coded 1 if a student's teacher obtained a degree other than bachelor and master degree (0 if not) | KCEWS |
| *School Characteristics* | |
| School previous year performance | CCD |
| School size | CCD |
| Student–teacher ratio | CCD |
| Percentage of White students in school | CCD |
| Percentage of Black students in school | CCD |
| Percentage of Hispanic students in school | CCD |
| Percentage of students in other racial groups in school | CCD |
| Percentage of students who are eligible for free or reduced-price lunch in school | CCD |
| Whether school is classified as "Distinguished" or "Distinguished/Progressing," coded 1 if a school is classified as "Distinguished" or "Distinguished/Progressing" (0 if not) | CCD |

| Variable Description | Source |
|---|---|
| Whether school is classified as "Proficient" or "Proficient/Progressing," coded 1 if a school is classified as "Proficient" or "Proficient/Progressing" (0 if not) | CCD |
| Whether school is classified as "Needs Improvement/Progressing," coded 1 if a school is classified as "Needs Improvement/Progressing" (0 if not) | CCD |
| Whether school is classified as "Needs Improvement," coded 1 if a school is classified as "Needs Improvement" (0 if not) | CCD |

*Note.* CCD = Common Core of Data; KCEWS = Kentucky Center for Education and Workforce Statistics.

Next, every student taught by NBCTs was matched to one taught by non-NBCTs on the estimated propensity scores. We implemented 1:1 nearest-neighbor matching without replacement. The matching was carried out using the MatchIt package in *R* (Ho, Imai, King, & Stuart, 2007).

## Checking Baseline Equivalence

The equivalence of the comparison groups of students was assessed by evaluating balance on key baseline measures (e.g., student's pretest score, race/ethnicity).[5] For the balance diagnostic test, the standardized difference between the treatment and control groups was divided by the pooled standard deviation for each measure. Following What Works Clearinghouse™ procedures (version 3.0; Institute of Education Sciences, 2014), when the difference between treatment and comparison groups on observable baseline characteristics is greater than 0.25 SDs, the treatment and comparison groups are judged to be not equivalent. When the difference in baseline characteristics is between 0.05 and 0.25 SDs, the analysis should include a statistical adjustment for the baseline characteristics. Differences of less than or equal to 0.05 SDs require no statistical adjustment. Table 3 shows that the differences between North Carolina treatment and comparison groups on baseline pretreatment characteristics are less than 0.05 SDs for all variables except the percentage of students taught by a teacher with an advanced degree at fourth grade (standard mean difference [SMD] =-.09) and the percentage of students taught by a licensed teacher (SMD = .11) at fifth grade. Table 4 shows that the differences between Kentucky treatment and comparison groups on baseline pretreatment characteristics are less than 0.05 SDs for all variables except the percentage of students with disabilities at fourth grade (SMD = .07) and the percentage of Hispanic students (SMD = .08).

**Table 3. Standardized Mean Differences Between Treatment and Comparison Schools for North Carolina**

| Baseline Covariates | Grade 4 | Grade 5 |
|---|---|---|
| Female | .01 | -.01 |
| Racial minority | .00 | .01 |
| FRPL status | -.01 | .01 |

[5] Binary dummy variables were created to convert the race/ethnicity categorical variable into binary variables (e.g., White, Black) during the propensity score matching process. To examine the balance on the baseline measures, standardized mean differences (effect sizes) were calculated. Hedges' *g* was used to calculate the absolute effect size (ES) difference for continuous variables (e.g., test scores) and the Cox index was used to calculate the ES difference for the binary variables (e.g., FRLP status).

| Baseline Covariates | Grade 4 | Grade 5 |
|---|---|---|
| Limited English proficiency status | .00 | .02 |
| Student learning disability status | .00 | .01 |
| Student attendance rate | .00 | -.01 |
| Math pretest | .00 | .00 |
| Reading pretest | .00 | -.01 |
| In-school or out-of-school suspension experience | .01 | .01 |
| Students taught by advanced degree teachers | -.09 | .03 |
| Students taught by experienced teachers | .00 | .00 |
| Students taught by licensed teachers | -.03 | .11 |
| School previous year performance | -.01 | -.02 |
| School size | -.01 | -.01 |
| Percentage of students who are eligible for FRPL in school | .00 | .02 |
| Student–teacher ratio | .01 | -.03 |
| Percentage of minority (i.e., non-White) students in school | .02 | -.02 |
| Percentage of advanced degree teachers | .01 | .00 |
| Percentage of experienced teachers | .00 | .00 |
| Percentage of licensed teachers | -.01 | .00 |

*Note.* FRPL = free or reduced-price lunch.

**Table 4. Standardized Mean Differences Between Treatment and Comparison Schools for Kentucky.**

| Baseline Covariates | Grade 4 | Grade 5 |
|---|---|---|
| Math pretest | .00 | -.01 |
| Reading pretest | .01 | .00 |
| In-school suspension experience | .01 | .02 |
| Out-of-school suspension experience | .03 | .01 |
| FRPL status | .00 | -.02 |
| Limited English proficiency | -.01 | .02 |
| Student learning disability status | -.07 | .00 |
| Female | -.01 | -.02 |
| White | .00 | -.04 |
| Black | -.01 | .02 |
| Hispanic | .01 | .08 |
| Students taught by National Board Certified Teachers | -.02 | .02 |
| School size | -.01 | .03 |
| School academic performance | .00 | .03 |

| Baseline Covariates | Grade 4 | Grade 5 |
|---|---|---|
| School ranking | .00 | .03 |
| Percentage of students who are eligible for FRPL | .00 | -.01 |
| Percentage of White students | .00 | .01 |
| Percentage of Black students | -.01 | -.02 |
| Percentage of Hispanic students | .01 | .00 |

*Note.* FRPL = free or reduced-price lunch.

## Outcome Measures

Each study contained the same academic and behavioral outcomes. There are two academic outcomes and three behavioral outcomes for each study.

### *North Carolina End-of-Grade (NC EOG) Mathematics and Reading Comprehension Test Scores (Grades 4 and 5)*

The NC EOG exam is administered to students in Grades 3–8. It is a curriculum-based multiple-choice achievement test aligned to the North Carolina Standard Course of Study.

### *Kentucky Performance Rating for Educational Progress (K-PREP) Mathematics and English Test Scores (Grades 4 and 5)*

The K-PREP exam is administered to students in Grades 3–8. It is a blended model built with norm-referenced and criterion-referenced test items that consist of multiple-choice, extended-response, and short-answer items that are aligned to the Kentucky Academic Standards.

### *Attendance Rate*

The attendance rate for each student was calculated by dividing the number of days enrolled by the number of days in attendance.

### *In-School Suspensions*

This is a binary outcome variable where students who received an in-school suspension are coded as 1 and all other are coded as 0.

### *Out-of-School Suspensions*

This is a binary outcome variable where students who received an out-of-school suspension are coded as 1 and all other are coded as 0.

Tables 5 and 6 contain average, unadjusted outcomes by grade level for North Carolina and Kentucky, respectively.

**Table 5. Descriptive Statistics of 2014–15 North Carolina Student Outcomes, by Grade and Treatment Status**

| Grade | Student Outcome Measure | Treatment | | | Comparison | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SD | *N* | Mean | SD | *N* |
| 4 | NC EOG mathematics score | 448.59 | 40.12 | 8,058 | 448.38 | 36.21 | 8,058 |
| | NC EOG reading score | 444.61 | 39.47 | | 444.7 | 35.7 | |
| | Attendance rate | 0.97 | 0.03 | | 0.97 | 0.03 | |
| | In-school suspension | 0.02 | 0.14 | | 0.02 | 0.14 | |
| | Out-of-school suspension | 0.03 | 0.18 | | 0.03 | 0.18 | |
| 5 | NC EOG mathematics score | 449.64 | 36.7 | 7,328 | 449.77 | 30.88 | 7,328 |
| | NC EOG reading score | 448.68 | 36.44 | | 448.85 | 30.63 | |
| | Attendance rate | 0.97 | 0.03 | | 0.97 | 0.03 | |
| | In-school suspension | 0.03 | 0.17 | | 0.02 | 0.15 | |
| | Out-of-school suspension | 0.04 | 0.2 | | 0.04 | 0.19 | |

*Note.* SD = standard deviation.

**Table 6. Descriptive Statistics of 2014–15 Kentucky Student Outcomes, by Grade and Treatment Status**

| Grade | Student Outcome Measure | Treatment | | | Comparison | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SD | *N* | Mean | SD | *N* |
| 4 | K-PREP mathematics score | 212.29 | 18.23 | 2,695 | 212.63 | 19.20 | 2,695 |
| | K-PREP reading score | 212.63 | 15.32 | | 212.36 | 15.34 | |
| | Attendance rate | 0.96 | 0.04 | | 0.96 | 0.04 | |
| | In-school suspension | 0.02 | 0.19 | | 0.03 | 0.36 | |
| | Out-of-school suspension | 0.01 | 0.12 | | 0.02 | 0.19 | |
| 5 | K-PREP mathematics score | 217.42 | 20.71 | 2,524 | 216.21 | 20.91 | 2,524 |
| | K-PREP reading score | 214.62 | 15.01 | | 213.68 | 15.25 | |
| | Attendance rate | 0.96 | 0.04 | | 0.96 | 0.04 | |
| | In-school suspension | 0.04 | 0.41 | | 0.04 | 0.39 | |
| | Out-of-school suspension | 0.02 | 0.17 | | 0.02 | 0.20 | |

*Note*. SD = standard deviation.

## Analytic Approach

The outcome analysis examines the effect of NBCTs on students' academic (e.g., reading and mathematics end-of-year test) (research question 1) and behavioral outcomes (attendance rate and suspension outcomes) (research question 2) during 2014-2015. The general modeling approach was performed separately for each of the four outcomes and for each grade level.

$$Y_i = \beta_0 + \beta_1(\text{NBCT})_i + \beta_2(\text{Prior Year Outcomes})_i + \beta_3(PS - \text{logit})_i + \boldsymbol{B}_4 X_i + \boldsymbol{B}_5 Z_j + e_i$$

where $Y_i$ is the outcome measure (i.e., standardized reading and mathematics end-of-grade test scores, attendance rate, or suspension[6]) for a student $i$. A treatment indicator for a student $i$ who was taught by an NBCT during 2014–15 ($\text{NBCT}_i$) is entered into the model after controlling for 2013–14 school year outcomes[7] ($\text{Prior Year Outcomes}_i$) and the logit of the estimated propensity score ($\text{PS-logit}_i$). The vector $\boldsymbol{X}$ includes student and teacher characteristics for student $i$.[8] To account for the nested data structure in the data and to eliminate bias in the estimate attributed to school differences, we included school fixed effects,[9] the vector $\mathbf{Z}$. This school fixed effects controls for many observable and unobservable school characteristics.

To examine the differential effect of NBCTs by selected characteristics (e.g., FRPL status), we augmented the equation above with an interaction term to indicate whether there is a statistically significant difference in the treatment effect by selected characteristics. $\text{Subgroup}_i$ is the indicator for whether student $i$ belongs to a subgroup (e.g., FRPL status).

$$Y_i = \beta_0 + \beta_1(\text{NBCT})_i + \beta_2(\text{Subgroup})_i i + \beta_3(\text{Subgroup})_i * (\text{NBCT})_i + \beta_4(\text{Prior Year Outcomes})_i + \beta_5(\text{PS} - \text{logit})_i + \boldsymbol{B}_6 X_j + \boldsymbol{B}_7 Z_i + e_i$$

# Results

The results are presented separately for each study given that each state represents independent samples. We first present the results from the North Carolina study followed by the Kentucky study. The results of the academic outcomes will be followed by the results of the behavioral outcomes. For each study, we present all results by separate grade levels—first for Grade 4 and

[6] A linear regression model was used for the continuous variables including the test scores and logit transformed attendance rate; a logistic regression model was used for the in-school and out-of-school suspension binary variables.

[7] The prior-year outcome helps remove residual within-school bias associated with the propensity score and improves precision of estimation (Hong & Yu, 2007).

[8] Although the propensity score matching was the primary method that was used to control for differences between treatment and comparison students, covariates that exceeded .05 SMD between the two groups were included as additional controls in the respective outcome model (see Tables 3 and 4).

[9] We choose school fixed effects rather than school random effects because in an observational study, the assumption of the random effects that unobserved differences among the intact group should be uncorrelated with other predictors in the model can be violated (Murnane & Willett, 2010). Even though treatment and comparison students were matched using key baseline school characteristics (e.g., school performance, school size, percentage FRPL, student–teacher ratio, etc.) within a district, some unmeasured school characteristics can be correlated with the treatment indicator.

followed by Grade 5 results. As independent samples, we treat each grade separately within each state to generate grade-specific results rather than pooling the estimates across grades.

## Study 1: North Carolina

### *Effects of Being Taught by NBCTs on Student Academic Outcomes*

Table 7 shows the results of Grade 4 and Grade 5 analyses that estimate the effects of being taught by NBCTs on student achievement in mathematics and reading. There were no statistically significant effects on the student mathematics and reading outcomes.

**Table 7. Estimates of the Effects of Being Taught by NBCTs on North Carolina Student Academic Outcomes, by Grade**

| Outcome | Grade 4 | | Grade 5 | |
|---|---|---|---|---|
| | Estimate | Effect Size | Estimate | Effect Size |
| Math | 0.176 | .002 | -0.018 | .000 |
| | (0.207) | | (0.179) | |
| Reading | -0.113 | -.001 | 0.011 | .000 |
| | (0.207) | | (0.185) | |

*Note*. Standard errors are presented in parentheses.
*$p < .05$.

### *Effects of Being Taught by NBCTs on Student Behavioral Outcomes*

Table 8 shows the results of Grade 4 and Grade 5 analyses that estimate the effects of being taught by NBCTs on student attendance and suspension outcomes. The effects on the Grade 5 attendance rate were statistically significant in that the attendance rate of students of NBCTs is 0.02 SD units higher than students of non-NBCTs ($p < .05$). There were no statistically significant effects on the likelihood of a student receiving an in-school or out-of-school suspension.

**Table 8. Estimates of the Effects of Being Taught by NBCTs on North Carolina Student Behavioral Outcomes, by Grade**

| Outcome | Grade 4 | Grade 5 |
|---|---|---|
| | Estimate | Estimate |
| Attendance rate | 0.039 | 0.111* |
| | (0.048) | (0.051) |
| In-school suspension | 0 | 0.005 |
| | -0.003 | (0.003) |
| Out-of-school suspension | 0.002 | -0.002 |
| | (0.003) | (0.004) |

*Note*. Standard errors are presented in parentheses. The attendance rate (AR) estimate was based on the logit transformed attendance rate using the formula: log (AR/(1 – AR)) because the original metric was a percentage that was bounded by 0 and 1.
*$p < .05$.

### *Student Subgroup Analyses*

Exploratory subgroup analyses were conducted to examine the differential effect of NBCTs by the following student subgroups: students with disabilities, students with limited English proficiency, students eligible for FRPL, sex, and race/ethnicity. Table 9 shows the results of these subgroup analyses. Four statistically significant differential effects were detected for two of the five subgroups. The Grade 5 reading difference and the Grade 4 and Grade 5 mathematics difference between students of NBCTs and students of non-NBCTs is larger for students without disabilities compared with the gap between students of NBCTs and students of non-NBCTs who do have disabilities (0.01 SD units for Grade 5 reading as well as Grade 4 and Grade 5 mathematics). Therefore, the effect of having an NBCT on Grade 5 reading, and Grade 4 and Grade 5 mathematics is larger for students who do not have disabilities than it is for students with disabilities. Finally, the Grade 5 attendance rate difference between students of NBCTs and students of non-NBCTs is larger for White students compared with the difference between students of NBCTs and students of non-NBCTs who are not White (0.02 SD units). Therefore, the effect of having an NBCT on Grade 5 attendance rates is larger for White students than it is for non-White students.

**Table 9. Estimates of North Carolina Subgroup Analyses**

| Outcome | SWD | | LEP | | FRPL | | Gender | | Minority | |
|---|---|---|---|---|---|---|---|---|---|---|
| | NBCT by SWD Interaction | | NBCT by LEP Interaction | | NBCT by FRPL Interaction | | NBCT by Female Interaction | | NBCT by Minority Status Interaction | |
| | Grade 4 | Grade 5 | Grade 4 | Grade 5 | Grade 4 | Grade 5 | Grade 4 | Grade 5 | Grade 4 | Grade 5 |
| Reading | -0.851 | -1.245* | 0.246 | -0.988 | 0.090 | 0.374 | 0.512 | -0.305 | 0.211 | 0.430 |
| | (0.598) | (0.510) | (0.608) | (0.785) | (0.361) | (0.321) | (0.346) | (0.306) | (0.368) | (0.327) |
| Math | -1.220* | -2.219*** | -0.116 | 0.578 | -0.050 | 0.468 | 0.448 | -0.324 | -0.048 | 0.180 |
| | (0.599) | (0.494) | (0.609) | (0.762) | (0.362) | (0.311) | (0.346) | (0.297) | (0.369) | (0.317) |
| Attendance rate | -0.196 | -0.037 | 0.022 | 0.089 | -0.075 | -0.058 | 0.023 | 0.023 | -0.144 | -0.187* |
| | (0.138) | (0.141) | (0.140) | (0.217) | (0.083) | (0.089) | (0.080) | (0.085) | (0.085) | (0.090) |
| In-school suspension | -0.005 | 0.009 | -0.013 | -0.006 | 0.003 | -0.005 | -0.004 | -0.005 | -0.003 | 0.001 |
| | (0.007) | (0.009) | (0.007) | (0.013) | (0.004) | (0.005) | (0.004) | (0.005) | (0.005) | (0.006) |
| Out-of-school suspension | 0.007 | -0.003 | -0.000 | -0.003 | 0.007 | -0.007 | 0.002 | 0.001 | 0.004 | 0.004 |
| | (0.009) | (0.010) | (0.009) | (0.016) | (0.006) | (0.006) | (0.005) | (0.006) | (0.006) | (0.007) |

*Note.* FRPL = free or reduced-price lunch; LEP = limited English proficiency; NBCT = National Board Certified Teacher; SWD = student with disability. The attendance rate (AR) estimate was based on the logit transformed attendance rate using the formula: log (AR/(1 – AR)) because the original metric was a percentage that was bounded by 0 and 1.
*$p < .05$. ***$p < .001$.

## Study 2: Kentucky

### *Effects of Being Taught by NBCTs on Student Academic Outcomes*

Table 10 shows the results of Grade 4 and Grade 5 analyses that estimate the effects of being taught by NBCTs on student achievement in mathematics and reading. The effects on Grade 5 mathematics and reading were statistically significant in that students of NBCTs scored 0.06 SD units higher than students of non-NBCTs on both outcomes ($p < .05$). However, no effects on Grade 4 mathematics or reading were detected.

**Table 10. Estimates of the Effects of Being Taught by NBCTs on Kentucky Student Academic Outcomes, by Grade**

| Outcome | Grade 4 | | Grade 5 | |
|---|---|---|---|---|
| | Estimate | Effect Size | Estimate | Effect Size |
| Math | -0.34 | -.02 | 1.21* | .06 |
| | (0.50) | | (0.58) | |
| Reading | 0.27 | .02 | 0.94* | .06 |
| | (0.41) | | (0.43) | |

*Note*. Standard errors are presented in parentheses.
*$p < .05$.

### *Effects of Being Taught by NBCTs on Student Behavioral Outcomes*

Table 11 shows the results of Grade 4 and Grade 5 analyses that estimate the effects of being taught by NBCTs on student attendance and suspension outcomes. There were no statistically significant effects on the student attendance rate or the likelihood of a student receiving an in-school or out-of-school suspension.

**Table 11. Estimates of the Effects of Being Taught by NBCTs on Kentucky Student Behavioral Outcomes, by Grade**

| Outcome | Grade 4 | Grade 5 |
|---|---|---|
| | Estimate | Estimate |
| Attendance rate | 0.05 | 0.24 |
| | (0.30) | (0.30) |
| In-school suspension | 0.03 | -0.19 |
| | (0.25) | (0.19) |
| Out-of-school suspension | -0.45 | -0.11 |
| | (0.30) | (0.25) |

*Note*. Standard errors are presented in parentheses. The attendance rate (AR) estimate was based on the logit transformed attendance rate using the formula: log (AR/(1 – AR)) because the original metric was a percentage that was bounded by 0 and 1.

### *Student Subgroup Analyses*

Exploratory, subgroup analyses were conducted to examine the differential effect of NBCTs by the following student subgroups: students with disabilities, students with limited English proficiency (LEP), students eligible for FRPL, sex, and race/ethnicity. Table 12 shows the results of these subgroup analyses. Five statistically significant differential effects were detected for four of the five subgroups. The Grade 4 mathematics difference between students of NBCTs and students of non-NBCTs is larger for students eligible for FRPL than the difference between students of NBCTs and students of non-NBCTs who are not eligible for FRPL (0.07 SD units). Therefore, the effect of having an NBCT on Grade 4 mathematics is larger for students eligible for FRPL than it is for students who are not eligible for FRPL. The Grade 5 mathematics difference between students of NBCTs and students of non-NBCTs is larger for students with disabilities compared with the gap between students of NBCTs and students of non-NBCTs who do not have disabilities (0.06 SD units). Therefore, the effect of having an NBCT on Grade 5 mathematics is larger for students with disabilities than it is for students who do not have disabilities. The Grade 5 attendance rate difference between students of NBCTs and students of non-NBCTs is larger for females than the difference between students of NBCTs and students of non-NBCTs who are males (0.07 SD units). Therefore, the effect of having an NBCT on attendance rates is larger for females than for males. The Grade 4 reading difference between students of NBCTs and students of non-NBCTs is larger for non-White students than the difference between students of NBCTs and students of non-NBCTs who are White (0.06 SD units). Therefore, the effect of having an NBCT on Grade 4 reading is larger for non-White students than it is for White students. Finally, the Grade 5 attendance rate difference between students of NBCTs and students of non-NBCTs is larger for non-White students than the difference between students of NBCTs and students of non-NBCTs who are White (0.06 SD units). Therefore, the effect of having an NBCT on Grade 5 attendance rates is larger for non-White students than it is for White students.

**Table 12. Estimates of Kentucky Subgroup Analyses**

| Outcome | SWD | | LEP | | FRPL | | Gender | | Race/Ethnicity | |
|---|---|---|---|---|---|---|---|---|---|---|
| | NBCT by SWD Interaction | | NBCT by LEP Interaction | | NBCT by FRPL Interaction | | NBCT by Female Interaction | | NBCT by White Interaction | |
| | Grade 4 | Grade 5 | Grade 4 | Grade 5 | Grade 4 | Grade 5 | Grade 4 | Grade 5 | Grade 4 | Grade 5 |
| Math | 2.11<br>(1.62) | 4.06*<br>(1.85) | 1.35<br>(3.07) | -3.71<br>(4.60) | 2.78**<br>(0.91) | -0.13<br>(1.06) | 0.37<br>(0.95) | 0.12<br>(1.08) | -2.06<br>(1.24) | 0.87<br>(1.37) |
| Reading | -0.45<br>(1.33) | 2.17<br>(1.35) | 4.62<br>(2.52) | -4.38<br>(3.35) | 0.90<br>(0.75) | 0.46<br>(0.78) | 0.53<br>(0.78) | 0.68<br>(0.79) | -2.48*<br>(1.02) | -0.13<br>(1.00) |
| Attendance rate | -0.50<br>(0.98) | 0.39<br>(0.99) | -0.22<br>(1.87) | 0.18<br>(2.47) | 0.77<br>(0.58) | -0.61<br>(0.59) | 1.06<br>(0.57) | 1.39 *<br>(0.58) | 0.61<br>(0.76) | -1.60*<br>(0.73) |
| In-school suspension | -0.23<br>(0.68) | -0.47<br>(0.54) | -1.41<br>(1.26) | -18.69[a]<br>(4086.23) | 0.50<br>(0.59) | 0.12<br>(0.39) | 0.26<br>(0.53) | -0.12<br>(0.45) | 1.01<br>(0.56) | -0.12<br>(0.40) |
| Out-of-school suspension | 0.59<br>(0.72) | -0.55<br>(0.65) | -18.93[a]<br>(5074.26) | -0.10[a]<br>(6138.28) | 0.60<br>(0.76) | -0.42<br>(0.53) | -1.73<br>(1.12) | -0.04<br>(0.66) | 0.45<br>(0.66) | -0.42<br>(0.55) |

*Note*. FRPL = free or reduced-price lunch; LEP = limited English proficiency; NBCT = National Board Certified Teacher; SWD = student with disability. The attendance rate (AR) estimate was based on the logit transformed attendance rate using the formula: log (AR/(1 – AR)) because the original metric was a percentage that was bounded by 0 and 1.

[a] The coefficient estimates and standard error are inflated due to the perfect separation between 0 and 1 in the outcome variable by the interaction term.

*$p$ < .05. **$p$ <.01.

# Discussion

This working paper reports the results from two studies that examined the effects of NBCTs on student academic and behavioral outcomes. Compared with previous studies that have had the same aim, the current work used more recent data and included additional outcomes not previously examined. Specifically, we assessed outcomes from the 2014–15 school year and examined both achievement and behavioral outcomes. The achievement outcomes are from Common Core–aligned mathematics and reading state assessments, and the behavioral outcomes included attendance rate and whether a student received an in-school or out-of-school suspension.

In this discussion, we first provide a brief summary of the North Carolina achievement results, followed by the Kentucky student achievement results. We do not intend for the results from this work to create a horserace between the two states. Thus, such results are discussed separately. In doing so, we add to the mixed findings related to the effects of NBCTs on student achievement given that both of the current studies generated findings that do not align perfectly with similar, past work. Then, we briefly discuss the behavioral outcomes results.

The North Carolina student achievement results are inconsistent with findings from similar studies that found positive effects of NBCTs on student achievement, including those conducted in North Carolina. In the four previously noted North Carolina studies that used student achievement data from between the 1996–97 and 2003–04 school years, those study authors all found some evidence that NBCTs had a positive effect on student achievement. It is possible that the lack of NBCT effects in the current North Carolina study could be attributed to this study's use of Common Core–aligned state assessments. However, that seems unlikely given that positive effects were detected in the Kentucky sample, for which we also used Common Core–aligned state assessments.

Whereas the North Carolina study did not generate student achievement findings that are consistent with past work, the Kentucky study results are more consistent with those from past studies. Specifically, Grade 5 students of NBCTs scored higher on state mathematics and reading assessments than students of non-NBCTs. This finding is noteworthy given that this is the first study of NBCT effects to use Kentucky data and replicates earlier findings from Florida, Washington, and North Carolina.

We noted that this study is the first to our knowledge that examined the effect of NBCTs on student behavioral outcomes. We detected positive effects of having an NBCT on student attendance in North Carolina. However, we did not detect any other positive effects, and we did not detect any negative effects. We speculate that there could be two reasons for limited NBCT effects on Grade 4 and Grade 5 attendance and no effects on the suspension outcomes. First, it is possible that any such effects are more likely to be explained by teacher experience (see, for example, Ladd & Sorensen, 2015). Second, it is possible that such effects are more likely to be detected when students are older (see, for example, Jackson, 2016).

Although this work, like other nonexperimental studies, adds to the evidence that NBCTs, on average, have an effect on student achievement, we believe that more experimental evidence is

needed to assess the impact of NBCTs on student outcomes. Such outcomes, in both future experimental and nonexperimental studies, should include not only student achievement on state assessments but other outcomes as well. For example, we know that teachers can have an effect on student outcomes such as collaboration skills, academic engagement, motivation to learn, and self-efficacy (Zeiser, Taylor, Rickles, Garet, & Segeritz, 2014). More work is needed to examine a fuller theory of action that explains the specific classroom practices that account for effects on student achievement and other outcomes. One way to accomplish this is for independent raters to observe NBCTs and non-NBCTs, and to examine whether specific classroom practices mediate the link between NBCT status and student outcomes. Given the existing evidence regarding the positive effects of NBCTs on student achievement, it would be useful to know which specific NBCT practices, exhibited more so than by non-NBCTs, may account for effects of NBCTs on student outcomes. Such information could be used by instructional coaches to deliver targeted support to teachers who do not demonstrate such practices. It could also be used by district and school leaders who are looking for the most effective teachers to fill their classrooms. In the absence of a presumed effectiveness signal, such as the NBCT credential, district and school leaders could consider identifying teachers who demonstrate NBCT-like practices to teach their students.

# References


Cantrell, S., Fullerton, J., Kane, T. J., & Staiger, D. O. (2008). *National Board certification and teacher effectiveness: Evidence from a random assignment experiment* (NBER Working Paper 14608). Cambridge, MA: National Bureau of Economic Research.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679.

Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review*, *30*(3), 449–465.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, *41*(4), 778–820.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*(6), 673–682.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (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.

Cowan, J., & Goldhaber, D. (2016). National Board certification and teacher effectiveness: Evidence from Washington state. *Journal of Research on Educational Effectiveness*, *9*(3), 233–258.

Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy, 11*(2), 125–149.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching. *Review of Economics and Statistics*, *89*(1), 134–150.

Hanushek, E. (2010). The difference is teacher quality. In K. Weber (Ed.), *Waiting for "Superman": How we can save America's failing public schools* (pp. 81–100). New York, NY: Public Affairs.

Harris, D. N., & Sass, T. R. (2009). The effects of NBPTS-certified teachers on student achievement. *Journal of Policy Analysis and Management*, *28*(1), 55–80.

Ho, D., Imai, K., King, G., & Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15,* 199–236.

Hong, G., & Yu, B. (2007). Early-grade retention and children's reading and math learning in elementary years. *Educational Evaluation and Policy Analysis*, *29*(4), 239–261.

Institute of Education Sciences. (2014). *What Works Clearinghouse: Procedures and standards handbook, Version 3.0*. Washington, DC: Author.

Jackson, K. C. (2016). What do test scores miss? The importance of teacher effects on non-test score outcomes (Working paper). Retrieved from, https://works.bepress.com/c_kirabo_jackson/30/

Ladd, H., & Sorensen, L. (2015). *Returns to teacher experience: Student achievement and motivation in middle school*. National Center for Analysis of Longitudinal Data in Education Research (CALDER). Washington, DC: CALDER. Retrieved from http://www.caldercenter.org/publications/returns-teacher-experience-student-achievement-and-motivation-middle-school

Murnane, R. J., & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research*. London, UK: Oxford University Press.

National Board for Professional Teaching Standards (NBPTS). (2015). *State incentive chart*. Arlington, VA: Author. Retrieved from http://www.nbpts.org/sites/default/files/Policy/state_incentive_chart.pdf

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*, 237–257.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 79*, 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, *94*(2), 247–252.

Zeiser, K. L., Taylor, J., Rickles, J., Garet, M. S., & Segeritz, M. (2014). *Evidence of deeper learning outcomes. Report #3 Findings from the Study of Deeper Learning: Opportunities and outcomes*. Washington, DC: American Institutes for Research. Retrieved from http://www.air.org/sites/default/files/downloads/report/Report_3_Evidence_of_Deeper_Learning_Outcomes.pdf

## ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

AIR®

AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

**www.air.org**

*Making Research Relevant*

## LOCATIONS

### Domestic

Washington, D.C.
Atlanta, GA
Austin, TX
Baltimore, MD
Cayce, SC
Chapel Hill, NC
Chicago, IL
Columbus, OH
Frederick, MD
Honolulu, HI
Indianapolis, IN
Metairie, LA
Naperville, IL
New York, NY
Rockville, MD
Sacramento, CA
San Mateo, CA
Waltham, MA

### International

Egypt
Honduras
Ivory Coast
Kyrgyzstan
Liberia
Tajikistan
Zambia

8625_02/17