Federal Reserve Bank of Chicago

# The Signaling, Screening, and Human Capital Effects of National Board Certification: Evidence from Chicago and Kentucky High Schools

*Lisa Barrow, Linda Cavalluzzo, Thomas Geraghty, Christine Mokher, and Lauren Sartain*

January 2020

WP 2020-06

https://doi.org/10.21033/wp-2020-06

# The Signaling, Screening, and Human Capital Effects of National Board Certification: Evidence from Chicago and Kentucky High Schools

Lisa Barrow, Federal Reserve Bank of Chicago; Linda Cavalluzzo, CNA; Thomas Geraghty, CNA; Christine Mokher, Florida State University; and Lauren Sartain, University of North Carolina at Chapel Hill

January 2020

The National Board for Professional Teaching Standards recognizes teachers who meet performance standards for "accomplished" educators. States and districts provide support for teachers to obtain this certification, which is considered an honor in the field. Using high school data from Chicago and Kentucky, we examine whether participation in the time- and resource-intensive certification process improves teacher productivity and, ultimately, if recognized teachers are of higher quality than their non-certified peers. We find the certification process itself did not increase teacher productivity. Further, we find mixed evidence on whether certified teachers are more effective at raising test scores than non-certified teachers.

Acknowledgements: We would like to thank the National Board for Professional Teaching Standards for providing background information, helping us to understand the certification process, and supplying data. This study also would not have been possible without the support and data provided by the Kentucky Department of Education, Chicago Public Schools, and the University of Chicago Consortium on School Research. We also thank Amanda McFarland and Kelly Vosters for analytic support, as well as seminar participants at the Federal Reserve Bank of New York/NYU Education Seminar Series, the Chicago Education Research Presentation series, various University of Chicago workshops, and the UChicago Consortium data group for thoughtful comments and suggestions. Authors are listed in alphabetical order. The views and opinions expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System. Any errors are ours.

## I. Introduction

One of the most important issues facing education policymakers in the United States is how to prepare students to be productive citizens in an increasingly competitive global economy. The international and national test scores of US teenagers suggest that there is room for improvement. For example, U.S. 15-year-old students perform poorly relative to their international peers in mathematics and science (Kastberg et al., 2016), and at the same time national trends in twelfth-grade test scores in mathematics have declined since 1998 (National Center for Education Statistics, 2015). Teachers are arguably the most important school-based input into student learning, and teacher quality varies considerably across and within schools (Aaronson, Barrow, & Sander, 2007; Goldhaber, 2002; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). As such, several policies at the national level have focused on improving teacher quality as a strategy for improving student outcomes. In order to staff every classroom with a high-quality teacher, school districts and principals must be able to identify, hire, and retain high quality teachers and/or improve existing teacher quality with professional development. One means for doing this may be through the National Board for Professional Teaching Standards (NBPTS) certification, which was established in part to recognize exceptional teachers. In this paper, we examine the signaling and screening value of NBPTS certification, as well as the potential for the rigorous, resource-intensive, and time-consuming certification process to improve the productivity of teachers.

One major challenge in identifying high-quality teachers is that traditional measures of teacher quality are largely unrelated to any easily observable teacher characteristics, such as highest level of education (Clotfelter, Ladd, & Vigdor, 2007; Goldhaber, 2007); years of teaching experience beyond the first two or three years (Clotfelter et al., 2007; Goldhaber, 2002; Rivkin et al., 2005); or indicators of ability such as selectivity of undergraduate institution or test scores

(Goldhaber, 2002; 2007; Harris & Sass, 2007; Kane, Rockoff, & Staiger, 2008). Therefore, while teachers are important to the learning process and good teachers improve the outcomes of students, it is ultimately difficult to pinpoint specific attributes that identify high-quality teachers.

The NBPTS was established to help professionalize the field of teaching by providing an accepted definition of what "accomplished" teaching is and recognizing teachers who do their jobs well. As such, one way for teachers to demonstrate their skill level and successes in the classroom is by earning certification from NBPTS. Since being established in 1987, hundreds of thousands of teachers have participated in the NBPTS application process, more than 120,000 teachers have become National Board (NB) certified (NBPTS, 2019), and large financial investments have been made to further develop the certification program. As a result, there is a great deal of interest in identifying and measuring the full value to education systems of encouraging teachers to become NB certified. If NB certification has no impact on teacher quality and it cannot distinguish between lower- and higher-quality teachers, then much time and money is being wasted.

To investigate the value of NB certification and the certification process itself, we utilize administrative data for teachers and students from the state of Kentucky and the Chicago Public Schools (CPS). Specifically, we seek to answer the following questions:

1. Is NB certification an effective signal of teacher quality?
2. Does the NB certification process effectively screen candidates?
3. Does the NB certification process improve teacher productivity?

We find mixed evidence on whether NB certification is an effective signal of teacher quality. Once we account for school effects, having a NB-certified teacher in $10^{th}$ or $11^{th}$ grade has a modest effect on student test scores, whereas we find no effect in $8^{th}$ or $9^{th}$ grade. We also find mixed results on the value of NB certification as a screening tool. That is, successful applicants are, at best, slightly more effective at improving student test scores than unsuccessful applicants. Finally,

we find no evidence that participating in the application process makes teachers more effective at improving student test scores. While test scores are only one outcome, given that time and resources are limited, these findings raise questions about whether teachers should be encouraged to pursue NB certification over other professional development opportunities.

We begin by outlining the potential role of NB in improving student learning and reviewing the relevant literature. Next, we describe the setting and the data sources followed by a description of the methods and findings from our analyses of student test scores. We conclude by summarizing the key findings, the limitations of this study, and the implications both for future research and for practice.

## II. The role of National Board in improving student learning

An original goal of NBPTS was to build an authentic assessment system that could reliably measure what experienced teachers should know and be able to do (Carnegie Task Force on Teaching as a Profession, 1986), resulting in a rigorous, multifaceted evaluation program. Applicants select from among 25 certificate areas, which are based on the age of the students taught (e.g., adolescence and young adulthood) and the subject area of instruction (e.g., Mathematics). To apply, teachers must assemble and submit a portfolio of specific materials including artifacts from their classroom instruction and student work, video of their classroom interactions with students, written reflections analyzing the instructional practice evident in the videos and student work, and provide evidence of their involvement in activities outside the classroom that benefit student learning. In addition, teachers must pass a computer-based assessment of their content and pedagogical knowledge in their specialty area (NBPTS, 2019).

During the period studied, the entire process could take from many months to nearly two years to complete, depending largely on when an applicant submitted her initial application.

Applicants submitted their application forms, fees, and proof of eligibility and began developing their portfolios sometime between February and December of the first year, and portfolio entries had to be completed by May of year two. Computer-based assessments were administered between March and June of the second year. In all, applicants had 16 months to complete all the requirements, with the caveat that no portfolio entry could be more than 12 months old. Applicants did not find out their certification status until November or December of year two.

In an evolution to the original process, teachers who did not pass all sections of the certification could reapply and resubmit materials for the section(s) they did not pass previously. The reapplication cycle was 1 year, as opposed to the initial 2-year application window. Once awarded, NB certification is valid for 10 years, at which point teachers must reapply if they are interested in maintaining their certification status although this is being replaced by a "Maintenance of Certification" process for certificates expiring after 2020.

Teachers who decide to apply for NB certification generally have many support options available to them. Until the 2013-14 application cycle, NBPTS offered a preparatory professional development program called "Take One!" that provided interested individuals with information about the certification standards and allowed teachers to submit one video portfolio entry for scoring prior to formally applying. Some districts and state departments of education, including the Kentucky Department of Education (KDE) and CPS, have central office staff members dedicated to helping teachers become NB certified, and many postsecondary schools of education offer programs to help teachers prepare for the rigors of NB certification. Teachers may also informally rely on colleagues to help them reflect on their practices and build their portfolio.

In Kentucky, the Kentucky Education Association offers professional learning opportunities for teachers interested in applying for certification or renewal. It also provides

training for educators who are interested in serving as mentors to NB candidates. Chicago teachers have at least two support options (one through the district and another through the teachers' union) for ongoing candidate support during the NB application process. These programs provide weekly or biweekly meetings for candidate teachers to come together to review and revise their portfolios, as well as counseling on whether or not the time-consuming process is a good fit for them.

We empirically test the main ways in which NB can improve the quality of classroom teaching. The first has been the subject of much academic research—that NB certification can serve as an indicator of teacher quality. This implies both that high-quality teachers apply for NB certification (the signaling effect) and that the NB certification process does a good job of screening applicants and awarding certification to the most qualified teachers (the screening effect). If certification is a good indicator of teacher quality, then principals and district administrators can use NB certification to inform their decisions about staffing and leadership roles for experienced teachers. Namely, given a large enough supply of NB–certified teachers, principals and school districts can improve average teacher quality by staffing a large number of teaching positions with NB teachers. Finally, NB certification might improve average teacher quality by changing and improving teachers' practices. In other words, perhaps the NB certification process itself develops teacher human capital, regardless of the application outcome.

## III. Literature review

The end goal of most education policy interventions is to improve student outcomes, and a primary mechanism for increasing student learning is to ensure that students are exposed to high-quality teaching. Many researchers have studied the question of whether NB certification is indeed a good measure of teacher quality both by examining the extent to which NB certified teachers are

high quality and the extent to which NB applicants who achieve certification are of higher quality than teachers who apply but fail to achieve certification. This second distinction is important because teachers self-select into the applicant pool. If only the highest quality teachers apply but certification is random, we would still find that NB certified teachers were higher quality than non-NB certified teachers. Additionally, any program that can also improve teaching quality among existing teachers is quite valuable as it can directly expand the pool of high quality teachers.[1] There has been much less research on the potential for the NB certification process to increase teacher human capital.

### a. Signaling framework

As noted above, if NB certification identifies high quality teachers, then principals and school districts could use NB certification to allocate resources and staff more effectively. For example, some principals might target NB–certified teachers in the hiring process or use NB status to make class assignment decisions. Teachers might use NB certification themselves to signal to principals and districts that they are high quality teachers and in turn use the certification to get more desirable teaching assignments or other more favorable treatment in the teacher labor market.

Several studies have found evidence that obtaining NB certification has modest signaling value, i.e., that teachers with NB certification are indeed of slightly higher quality than teachers who are not certified (Cantrell, Fullerton, Kane, & Staiger, 2008; Cavalluzzo, 2004; Clotfelter, Ladd, & Vigdor, 2007; Cowan & Goldhaber, 2016; Goldhaber & Anthony, 2007; Horoi & Bhai, 2018). Most studies that identify the signaling effect of NB certification compare NB teachers and noncertified teachers, making statistical adjustments to account for the fact that teachers who

[1] Having a program such as NB certification may also have indirect effects on the number of high-quality teachers by encouraging teachers to enter or stay in the profession. Analysis of any such effects is beyond the scope of this paper.

participate in the NB certification process might otherwise differ from those who do not. The effect sizes estimated are generally small but statistically significant for reading and more mixed for mathematics (See McCaffrey and Rivkin (2007) for a review.).

**b. Screening framework**

Research also suggests that the NB certification process is good at distinguishing more-effective from less-effective teachers who apply for certification. In other words, NB–certified teachers are more effective, as measured by student achievement, than are applicants who complete the application process but do not achieve certification (Cavalluzzo, 2004; Clotfelter et al., 2007; Goldhaber & Anthony, 2007; Sanders et al., 2005). In general, these studies find that students taught by NB–certified teachers make statistically significant, larger test score gains than those taught by teachers who applied but were not certified. Effect sizes tend to be larger for math than for reading (Elliott, Koenig, & Hakel, 2008).

**c. Human capital framework**

Finally, much less research has focused on whether the NB certification process itself can improve an applicant's teaching ability. An individual's human capital stock can potentially be developed through investment in education, training, and professional development activities; and in the case of teachers one would hope that this investment would lead to better outcomes (greater learning, for example) for their students. Overall, though, the literature shows little to no effect of most professional development programs on student outcomes (e.g., Harris & Sass, 2007; Jacob & Lefgren, 2004; Podgursky, Springer, & Hutton, 2010). In particular, much of the funding for professional development is spent on "one-shot" workshops or other events not shown to translate into improvements in student outcomes (Garet, Porter, Desimone, Birman, & Yoon, 2011).

However, as with any educational intervention, the quality of professional development varies, from good to bad and everything in between. Research on professional development in Chicago Public Schools suggests that teachers benefit most from training that promotes ambitious, intellectually challenging instruction; occurs frequently and over time; exposes the teacher to content in his or her subject area; and features developments in pedagogical techniques (Smylie, Allensworth, Greenberg, Harris, & Luppescu, 2001). The U.S. Department of Education defines high-quality professional development as sustained and content focused, aligned with state learning standards, and focused on developing an understanding of "scientifically proven" instructional techniques (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Although the original motivation for NBPTS was not to build a strong professional development program, the certification process has the markings of one. The NB application process itself is sustained over time, and the application materials include a portfolio of lessons, assessments, and reflections prepared by the teacher and based on the students in his or her actual classroom. As a result, one might reasonably expect that participation in the NB certification process could improve a teacher's instruction and that better instruction would translate into better student outcomes.

Our primary question of interest, therefore, is whether participation in the NB certification process itself improves a teacher's effectiveness, regardless of whether or not the applicant achieves certification. In other words, is the NB certification process effective professional development? The extant literature leaves understudied, and unresolved, whether NB certification is more than a good signal of and screen for identifying high-quality teachers. Studies that try to capture human capital effects of NB certification compare different teachers who are at different stages in the certification process (before applying, applying, and after applying). They typically find that teachers' effectiveness declines marginally while they are applying, which could be a

result of their spending so much time and energy on their portfolio that it distracts from their teaching (Clotfelder et al., 2006; 2007; Goldhaber & Anthony, 2007; Harris & Sass, 2006; McCaffrey & Rivkin, 2007). These same studies produce mixed results about gains in teacher effectiveness after the application process ends. The primary limitations of the existing human capital research are that any observable gains in student learning might simply be due to certified teachers being better able to signal and sort into schools or to getting different teaching assignments after being certified. Namely, gains could be a function of certified teachers now teaching higher achieving students or in higher achieving schools. We propose a different approach to estimating the human capital effect: comparing an individual teacher to his or herself over time using a teacher fixed effects strategy. Although this approach has had limited use in the research literature (e.g., Harris & Sass, 2006), it should result in more accurate estimates of the ability of the NB certification process to increase teacher human capital.

## IV. **Data**

In order to study NB certification and in the human capital effects on teachers, we need large numbers of teachers entering the program. The state of Kentucky and the city of Chicago offer two such sites. Kentucky is an ideal state for this study. First, NB enjoys strong support there. Through the efforts of teachers and the financial support of the Teachers' National Certification Incentive Trust Fund, the state has become one of the largest producers of NB teacherss: in 2013/14, 2,925 or about 7 percent of the teaching workforce in Kentucky was NB certified.[2] This compares favorably with the national average of about 2 percent. To our knowledge, however, there has been no notable research on the effectiveness of NB teachers compared with noncertified teachers in the state. Kentucky is also appealing because while it is largely rural, it also has

[2] Calculated based on data provided by NBPTS.

suburban and urban centers, including the Louisville/Jefferson County metro area, with a 2010 population of about 750,000.[3] Furthermore, Kentucky uses ACT's Educational Planning and Assessment System (EPAS) to monitor growth in student achievement over time, a test that is also administered in many other schools across the nation. The state has a longitudinal data system that uses unique identifiers to track students across the state and over time, and can be used to link students to their teachers, to the courses they enroll in, and to their statewide assessments.

As a large urban district, CPS provide an excellent complement to the Kentucky sample. The city of Chicago has a population of 2.8 million, and its school system is home to more than 1,700 NB teachers, or roughly one-third of all NB teachers in the state of Illinois.[4] Like other large urban districts, CPS is racially and ethnically diverse. Like Kentucky, CPS also administered the EPAS for several years, and the district has a longitudinal data system that enables students to be tracked over time and linked to their teachers, courses, and test scores at the high school level.

### a. Data sources

Our analysis relies on administrative data from all public middle and high schools in the state of Kentucky and all public non-charter high schools in CPS linked to NB application data. Student-level data files were provided the Kentucky Department of Education and by CPS through the UChicago Consortium on School Research. These data files include school enrollment information, student course reports linked to the teacher of record for the course, test scores, and student demographic characteristics. In both locations, we have four years of data, allowing us to measure changes in student outcomes over time for two or three cohorts of students in Kentucky (depending on the test score outcome) and three cohorts in Chicago. In Kentucky, data are

[3] Data from the U.S. Census (www.census.gov).
[4] Authors' calculation based on CPS (2014) and National Board Resource Center (2012).

available for school years (SYs) 2007/08 through 2010/11; in Chicago, data are available for 2008/09 through 2011/12. Application data from NBPTS covers all applicants in Kentucky or Chicago who were first-time applicants from the 2000-2001 through the 2012-2013 cohorts.

*i. Student test scores*

At the time, both CPS and Kentucky used EPAS, which consists of three tests: EXPLORE®, PLAN®, and ACT®. According to ACT, Inc., the tests are aligned so that the score of the next test in the series can be predicted based on the prior test score. Each test results in five sub-area scores: English, mathematics, reading, science, and writing. The composite score is the average of all of the sub-area scores except for writing. For the analysis, we use subject-level test scores for English, mathematics, and science. The EXPLORE is administered in the fall of grade 8 in Kentucky and the fall of grade 9 in CPS. In both locations, the PLAN is administered in the fall of grade 10; and the ACT is administered in the spring of grade 11. One challenge to estimating the impacts of teachers on these student test scores is that many students will have multiple teachers between two test administrations making it difficult to attribute test score growth to a particular teacher. This is particularly challenging in Kentucky where students may have multiple teachers for the same subject over the course of a single academic year.

One advantage of using a nationally normed test series such as EPAS is that we can directly compare students in Kentucky and CPS to each other and to students across the country. The average ACT score in the Kentucky analysis sample is 18.8 compared with an average ACT score of 17.5 in the CPS analysis sample and 21.0 nationally in 2010. We standardized the scale scores for each subject test by subtracting the national mean score on the corresponding test from the student's test score, and then dividing by the national standard deviation. The results can be interpreted as effect sizes in standard deviation units. The standardization also allows the

magnitude of the effects to be directly compared across subject areas, test (EXPLORE, PLAN, ACT), and locales (CPS, Kentucky). Results are examined separately for English, math, and science. We also examine the results for the three subjects pooled.[5]

All models are value-added models (described below), so we control for each student's prior test score in order to attribute growth between the two test points to a teacher. We conduct two sets of analyses for this study: the first uses the EXPLORE as a pretest and the PLAN score as the outcome measure; the second analysis uses the PLAN as a pretest and the ACT score as the outcome. The analysis sample includes only students who have both pretest and posttest scores.[6]

ii. Student information

Both CPS and Kentucky administrative data on students include basic demographic information, such as gender and race/ethnicity, as well as socioeconomic status (based on free/reduced-price lunch (FRPL) eligibility) and special education status (students with Individualized Education Programs). Date of birth was used to calculate each student's age at the beginning of each school year. In addition, Kentucky has an indicator for English as a Second Language (ESL) status, and the number of days the student was absent during the school year.

The analytic sample in Chicago includes 69,741 students for the PLAN analysis and 48,546 for the ACT analysis. In Kentucky, the sample sizes are 80,490 for the PLAN and 114,465 for the ACT. (Some 34,903 Kentucky students are in both the PLAN and the ACT samples.)

In Table 1 we present average characteristics for students in our analysis sample for three mutually exclusive subgroups—students who had at least one teacher who ever achieved NB

[5] We did not examine test scores in reading or writing because those topics do not align to a specific teacher.

[6] The majority of students took each test one time; however, if a student has more than one test score, we use the score from the date of the earliest test, so the results are comparable to students who took the test only once.

certification, students who had at least one teacher who applied by did not achieve NB certification (and no successful NB applicant teachers), and those who had only teachers who had never applied for NB certification. The top panel presents these characteristic means for the students based on their 9th grade teachers (in CPS) or their 8th and 9th grade teachers (in Kentucky) while the bottom panel presents characteristic means for students based on their 10th and 11th grade teachers. In all cases, we see that students who spend at least one semester with a NB certified teacher are positively selected. In particular, CPS students with a NB certified teacher score 0.1 standard deviations above the national average on the EXPLORE compared with 0.4 standard deviations below the national average for students who have at least one teacher who failed to achieve NB certification and 0.5 standard deviations below the national average for students who have only teachers who have never applied for NB certification. In addition, compared with teachers who have not applied for NB certification, CPS NB certified teachers are in classrooms with students who are less likely to be Black, more likely to be white, less likely to be eligible for free or reduced-price school lunch, and less likely to be classified as special education students. In Kentucky, NB certified teachers are also less likely to have free or reduced-price school lunch or special education students than non-applicant teachers. However, in Kentucky NB certified teachers are more likely to have Black students and less likely to have White students. This is likely due to a greater number of NB certified teachers in more urban areas where there is greater ethnic diversity. As a result, we control for student demographic characteristics and include prior test score observations to attempt to control for the fact that students are not randomly assigned to teachers and that NB certification may even change the type of students in a teacher's classroom. To the extent that prior test scores and student characteristics do not fully account for strategic assignments of students and teachers, our estimates may be biased. We return to this point when we discuss the results.

iii. Teacher information

NBPTS provided certification application data for teachers in CPS and Kentucky starting with the 2000 applicant cohort and ending with the 2012 applicant cohort. These data include application date(s), number of times applied, and the final status of each application – achieved, not certified, or applicant withdrew application – for teachers of all subjects and grade levels. We also have information about the subject area and age category for certification. Over this 13-year period, there were 4,658 unique applicants from CPS, and 44 percent of them achieved NB certification. From Kentucky there were 4,746 unique applicants, and 54 percent of them achieved NB certification. Most applicants applied one time (71 percent for CPS, 67 percent for Kentucky); about 1 percent of teachers applied more than three times.

There is no unique teacher identification number that can be used to link the NBPTS data with the teacher records in the administrative data files from CPS or KDE. Instead, we matched the records using teachers' first names, last names, and email addresses. We started by identifying any exact matches in either address or first and last name in both files. Then we looked for cases where the names were similar but not exact. We manually checked these records and compared other characteristics in the two files, such as school name and subject area, to determine whether the records appeared to belong to the same person. The match rate is expected to be less than 100 percent because the administrative data files include only public school teachers, while the NBPTS data include other applicants such as administrators and private school teachers. For the years of our analysis, the match rate is 83 percent in Kentucky and 78 percent in Chicago.

In order to link students to their teachers, we used transcript files that account for all the courses in which a student enrolls and the teachers of each course. For both CPS and Kentucky, we include only teachers of core English, mathematics, and science courses in the analysis. If the

student took both a core course and an elective course in a particular subject area, we included the record from the core course in the analysis and included an indicator variable to reflect that the student was also enrolled in an elective course in the same subject area.

Students who have more than one core course in the same subject area taught by more than one teacher were flagged as having multiple teachers as were students who had only one core course that was taught by multiple teachers. Conversely, students without any courses in the core subject area were flagged as having no teachers. While we cannot identify the individual teacher responsible for teaching these students in those particular semesters/years, we do not want to drop them from the analytic dataset.

For CPS, we also have access to teacher personnel data which we use to assess how teacher characteristics differ between teachers who apply for NB certification (both those who are successful and those who are not) and teachers who never apply for NB certification. We present these characteristics by NB applicant status in Table 2. Compared with both applicants who are not successful and non-applicant teachers, NB certified teachers are less likely to be Black and more likely to be white. NB certified teachers are also more likely to hold a master's degree. Among applicant teachers, those who successfully achieve NB certification have fewer years of experience teaching in CPS on average compared with those who do not achieve NB certification, and male applicants are somewhat less likely to achieve than female applicants.

iv. School information

Most of the school-level data for Kentucky come from the Common Core of Data housed at the U.S. Department of Education's National Center for Education Statistics. These data provide publicly available characteristics about each school across the country and can be aggregated up to the district, state, or national level. Covariates include school size, student-teacher ratio, student-

administrator ratio (district level), percent Black students, percent Hispanic students, percent FRPL students, per pupil spending (district level), and school locale. For CPS, we calculate these school-level variables from the student-level data. We also use EPAS data from CPS and KDE to calculate school-level average scores on ACT, PLAN, and EXPLORE in each subject area.

## V. Methods

Because students are not randomly assigned to teachers, we take a value-added approach for analyzing the impacts of having a NB teacher by controlling for prior subject-level test scores in addition to student demographic characteristics. In this way we are comparing teacher impacts on student test score outcomes for teachers who serve similar students in terms of their prior test scores and demographics. To address each of the research questions, we compare different groups of teachers. We explore the first question, which asks whether NB certification is a good signal of teacher effectiveness, by comparing the effectiveness of NB teachers with teachers who are not certified. The second question, which considers the effectiveness of NB certification as a screening process, is answered by comparing teachers who apply for and achieve certification with those who apply for but do not achieve it. The third question addresses the professional developmental properties of the NB certification process itself by comparing the effectiveness of individual teachers against themselves at different stages (before, during, after) in their application process using a teacher fixed effects approach.

For each of the three research questions, the outcome variable is a student's English, mathematics, or science test score. One set of models uses the student's ACT scores as the outcome, with the student's previous PLAN scores as the prior test scores. These analyses capture the effect of having a NB teacher in grades 10 or 11 in Chicago and in Kentucky. A second set of models uses the student's PLAN scores as the outcome, with the student's previous EXPLORE

scores as the prior test scores. In this case, these analyses capture the effect of having a NB teacher in grade 9 in Chicago and grades 8 or 9 in Kentucky. Separate models are run for each subject – math, English, and science – to compare the NB effects across subject areas. We also run a combined model that includes all three subjects, with additional variables to indicate whether the observation outcome represents a math, English, or science test score. Results are also presented separately for Kentucky and CPS to compare the NB effects across contexts.

One difference between our study and other studies in this literature is that students may be exposed to instruction from multiple teachers in the same subject area between when the tests are administered. In Kentucky, students take the EXPLORE at the beginning of 8th grade, the PLAN at the beginning of 10th grade, and the ACT at the end of 11th grade. In Chicago, the testing calendar is the same for $10^{th}$ and $11^{th}$ grade while the EXPLORE is administered at the beginning of $9^{th}$ grade. Thus, depending on the analysis, the prior test score occurs three to four semesters before the outcome test score. Because there are multiple semesters between the prior score and the outcome, each student-level observation may involve more than one teacher. In Kentucky, we observe the student's course-taking each semester; so, for a given subject, approximately one quarter of students had two or more teachers between the outcome test score and the prior test score. In Chicago, core courses generally run for a full year; because we only observe student course taking on a year-by-year basis, rather than each semester, there will typically be one teacher (per student, per subject) between when the student takes the PLAN test and when she takes the EXPLORE test and at most two teachers (per student, per subject) between administration of the ACT and the PLAN.

As noted earlier, one challenge in estimating teacher effectiveness using longitudinal data systems, as we do here, is that neither teachers nor students are randomly assigned to their

classrooms, or to their schools. Education-minded parents choose housing taking school quality into account; teachers choose where to work based in part on the school's quality; the most effective school leaders find ways to recruit the best candidates; and once in their schools, principals assign students to teachers thoughtfully, not at random. As a result, there likely are systematic differences in student and teaching assignments that affect test scores, but that have nothing to do with NB certification. Because of this challenge, for each analysis we use a variety of statistical controls and present three different regression models to get a fuller picture of the likely true effect of NB certification on student test scores. Importantly, because teachers are not randomly assigned to participate in NBPTS, we additionally make use of a teacher fixed effect strategy when evaluating the human capital effects of the NBPTS certification program. In this way, we can compare a teacher's effectiveness during and after participating in the certification program to her effectiveness before applying for NB certification.

**a. Signaling effect**

To test for a signaling effect of NB certification, we compare the test scores of students who had one or more semesters with a NB teacher between the baseline and outcome tests with scores of students who had no NB teachers between the tests. If NB certification is an effective signal of teaching quality, then students taught by certified teachers should perform better on the outcome test than students taught by non-certified teachers after controlling for baseline test scores and other student characteristics. To estimate the signaling value of NB, we model student subject-level test-score outcomes using ordinary least squares (OLS) regressions as follows:

$$(1) \qquad TS_{is,t+1} = \beta_0 + \beta_1 TS_{ist} + \delta NB_{ist} + \boldsymbol{X'_{it}}\gamma + \boldsymbol{T'_{ijst}}\pi + \phi_s + \phi_t + \varepsilon_{ist}$$

where $TS_{is,t+1}$ represents outcome test score *TS* for student *i*, in subject *s* (mathematics, English, or science), in time period *t+1*; $TS_{ist}$ is the student's baseline test score in that same subject; $NB_{ist}$

is an indicator variable equal to 1 if the student had at least one NB certified teacher in subject *s* in any semester between the baseline and outcome test scores and 0 if not; $X_{it}$ represents a vector of student demographic characteristics (race/ethnicity, gender, eligibility for free/reduced-price lunch, and special education status), an indicator equal to one if the student is taking any elective courses in that subject, and an indicator equal to one if the student has multiple core course teachers in that subject; $T_{iist}$ is a vector of teacher *j*'s years of experience and experience squared[7]; $\phi_s$ are test subject fixed effects; and $\phi_t$ are year fixed effects. In some specifications we also control for school fixed effects or school characteristics (percent of students of different race/ethnicity categories, percent of students eligible for FRPL, percent special education students, the natural logarithm of total enrollment, and school-level average composite ACT scores) and the average student baseline test score corresponding to the student's teacher.[8] $\beta_0$, $\beta_1$, $\delta$, $\gamma$, and $\pi$ represent coefficients and coefficient vectors to be estimated, and $\varepsilon_{ist}$ is the error term. Our coefficient of interest, $\delta$, represents the average impact on student test scores in standard deviation units of having at least one NB certified teacher. We believe that controlling for baseline test scores at both the individual and teacher-level help account for unobservable characteristics that influence both the outcome variable and the assignment of students to teachers with and without NB certification.

### *b. Screening effect*

[7] Experience is years of experience in CPS or years that the teacher appears in the data in Kentucky (our best proxy for years of experience). For specifications in which there is only one year between the baseline and outcome tests (fall grade 9 and fall grade 10, respectively), years of experience corresponds to the grade 9 teacher. For specifications in which there are multiple academic years between baseline and outcome test scores, years of experience corresponds to the outcome year teacher.

[8] Again, for specifications in which there is only one year between the baseline and outcome tests, the teacher is the grade 9 teacher. For specifications with multiple academic years between baseline and outcome test scores, the teacher is the outcome year teacher.

To test for a screening effect, we compare the performance of students who had teachers who will ever achieve certification ("ever certified") with the performance of students who had teachers who have or will apply but not achieve NB certification ("never certified"). If the NB certification process is an effective screening device for high-quality teachers, then students taught by "ever certified" teachers should perform better on tests than students taught by "never certified" teachers, controlling for any differences in students assigned to ever and never NB certified teachers. To estimate whether the NB application process is an effective screen for teacher quality (i.e., does NB certify the "right" candidates?), we estimate the following equation:

$$(2) \qquad TS_{is,t+1} = \beta_0 + \beta_1 TS_{ist} + \delta_1 ever_{ist} + \delta_2 never_{ist} + \boldsymbol{X'_{it}}\gamma + \boldsymbol{T'_{ijst}}\pi + \phi_s + \phi_t + \varepsilon_{ist}$$

in which all variables are as described for equation (1) but we have changed the control variables of interest to be counts of the number of semesters between the baseline and outcome test score observations that student *i* had with a teacher in subject *s* who is or will be NB certified, $ever_{ist}$, and counts of the number of semesters between the baseline and outcome test score observations that student *i* had with a teacher in subject *s* who has or will apply but not achieve NB certification, $never_{ist}$.[9] Again, some specifications include school fixed effects or school characteristics and average baseline achievement of the students in a teacher's classroom; $\beta_0$, $\beta_1$, $\delta$, $\gamma$, and $\pi$ represent coefficients and coefficient vectors to be estimated; and $\varepsilon_{ist}$ is the error term.. Our coefficient of interest, $\delta_1$, represents the average impact on student test scores in standard deviation units of having one semester of instruction from a teacher who ever receives NB certification compared to having only non-applicant teachers. We are also interested in $\delta_2$, which is the average impact on student test scores in standard deviation units of having one semester of

[9] We also control for the number of semesters with a teacher for whom we do not know their achievement status.

instruction from a teacher who applies but never receives NB certification compared to having only non-applicant teachers. Post-estimation we test if $\delta_1$ equals $\delta_2$ to compare the estimated effects of having one semester with a successful applicant teachers to one semester with an applicant teacher who is not certified, On average, grade 9 students in Chicago have 0.18 semesters with an ever-certified teacher, and students in the grades 10/11 sample have 0.41 semesters with an ever-certified teacher. In Kentucky, students in the grade 8/9 sample have 0.08 semesters with an ever-certified teacher, while students in the grade 10/11 sample have 0.14 semesters with an ever-certified teacher.

### *c. Human capital effect*

To estimate the effect of the certification process itself on student achievement, we use a teacher fixed effects framework. These models estimate the effect of the NB application process on teacher effectiveness by comparing teachers to themselves as they move in and out of the application process. Specifically, we compare the student performance of teachers after they complete the application process ("past applicant") to the student performance of these same teachers when they were applicants ("current applicant") to the performance of their students before they start the certification process ("future applicant"). If the NB certification process itself is effective professional development, then we should expect to see a positive coefficient on the "past applicant" indicator—implying that, on average, a given teacher's students show more test score growth after the teacher has completed the NB process than they did before the teacher participated in NB. Additionally, some previous studies have found evidence that current applicants may be less effective than either past or future applicants. We can use this model to investigate any such potential effects in our sample. The human capital models are estimated using the same value-added specification as in the signaling and screening models, but the addition of

the teacher fixed effects alleviates some of the concerns of selection bias – in terms of selection by teachers into the NB application process as well as non-random assignment of students to teachers. We note that non-random assignment of students to teachers remains a potential source of bias if the assignment mechanism changes after a teacher applies to become a NB teacher.

We examine whether there is a human capital effect of the NB certification process by estimating equation (3).

(3) $TS_{is,t+1} = \beta_0 + \beta_1 TS_{ist} + \delta_1 Current_{jt} + \delta_2 Past_{jt} + \boldsymbol{X}'_{it}\gamma + \phi_j + \phi_s + \phi_t + \varepsilon_{ist}$

In this case, $\delta_1$ and $\delta_2$ are the coefficients of interest; $\delta_1$ is the parameter that captures the effect of being a current applicant versus being a future applicant, while $\delta_2$ compares past applicants to future applicants. Model (3) includes a set of teacher fixed effects, $\phi_j$, so identification of $\delta_1$ and $\delta_2$ come from comparing a teacher to herself over time as she moves through the application process.[10] Again, all specifications include student demographic characteristics, an indicator equal to one if the student is taking any elective courses in that subject, and subject and year fixed effects, some specifications include school fixed effects or school characteristics and average baseline achievement of the students in a teacher's classroom; $\beta_0$, $\beta_1$, $\gamma$, and $\pi$ represent coefficients and coefficient vectors to be estimated; and $\varepsilon_{ist}$ is the error term.

The application process has the potential to stretch over two academic years. Despite the ability to open an application in January in application year 1, the majority of applicants begin an application before the summer of application year 1. Because teachers generally wait until the summer to begin their applications, we code teachers as current applicants during one academic school year. Applicants who completed the process before 2008-09 are always coded as past

[10] These specifications also include an indicator equal to one if the teacher information is missing and an indicator equal to one if the student has more than one core course teacher.

applicants, and applicants who begin the process after 2011-12 are always coded as future applicants. Teachers are considered currently applying during a single school year. It is the within-teacher movement across application status that allows us to estimate the professional development, or human capital, effect of the NB application process.

## VI. Results

This section presents estimates of the relationships between NB status and student test scores for the signaling, screening, and human capital models. For each analysis, we present results for three different value-added specifications. In our most basic specification, we control for the baseline student test score, student characteristics, teacher experience, test subject fixed effects, and year fixed effects. In our second specification we additionally control for school characteristics and the average baseline student test score (at the teacher level), and in the final specification we substitute school fixed effects for school characteristics. We compare the results across specifications, grades, and sites and discuss how we think the differences and similarities inform our understanding of NB certification and the full value of participating in the application process.

### *a. Signaling effect*

In Table 3 we present estimates of the relationship between student test scores and having at least one NB teacher since the baseline test. The top panel presents these estimates for students who have a NB teacher during $9^{th}$ grade in Chicago and during $8^{th}$ and/or $9^{th}$ grade in Kentucky; the bottom panel presents these estimates for students who have a NB teacher during $10^{th}$ and/or $11^{th}$ grade. Left-hand quadrants correspond to estimates from Kentucky and right-hand quadrants correspond to estimates from Chicago. The coefficient on the NB teacher indicator can be interpreted as the effect in standard deviation units of having at least one NB teacher in that subject, taking into account the student's prior achievement among other covariates.

For all subjects combined we estimate that having at least one NB teacher is associated with higher test scores when we control for baseline test score, student-level characteristics, and teacher experience. In most cases, once we control for average incoming student test scores (at the teacher level) and school characteristics or school fixed effects, our coefficient estimates fall by more than half. This is consistent with our expectation that NB teachers are not randomly distributed across schools. For students having a NB teacher in 8th and/or 9th grade, our estimates from the pooled sample and most of the subject-specific regressions are no longer statistically different from zero. For Kentucky, we find that having a NB math teacher in 8th and/or 9th grade is associated with 7 percent of a standard deviation higher test scores.

The story is somewhat different for having a NB teacher in 10th or 11th grade where our estimated effects sizes are larger—0.038 in the Kentucky pooled sample and 0.048 in the Chicago pooled sample—and remain statistically different from zero. In the subject-specific specifications we continue to find no relationship between having an NB teacher in science class and ACT science test scores.

### *b. Screening effect*

To estimate the screening effect, we compare test scores for students of teachers who currently hold or in the future will hold NB certification with test scores for students of teachers who have applied for certification in the past, or will do so in the future, but who do not achieve certification. The omitted group consists of nonparticipating teachers.[11] We measure the screening effect by the difference between the coefficient on the status indicator for number of semesters with an "ever certified" teacher and the coefficient on the status indicator for number of semesters with a "never certified" teacher. We present these estimates in Table 4 which has the same layout

[11] We include an indicator variable for teachers who apply but for whom we do not observe their certification status.

as described for Table 3. Now, the interpretation of the estimate presented is the increase in test scores associated with having one semester with a NB teacher rather than a NB applicant who fails to achieve certification.

Overall, we find little evidence that test scores are higher for students who have an 8th or 9th grade teacher who is NB certified than for students who have an 8th or 9th grade teacher who applies for NB certification but fails to achieve, once we control for average incoming test scores at the teacher level and school characteristics or school fixed-effects. In Kentucky, this pooled result masks a marginally significant coefficient estimate for the math test of 0.053 standard deviations. The English results for CPS suggest that NB teachers may be more effective than those who fail to achieve, but the difference is only 0.025 standard deviations. For students having NB applicant teachers in 10th or 11th grade, our pooled estimate is that certified teachers are associated with 0.02 standard deviation higher student test scores than teachers who fail to achieve. In this case, the subject-specific estimates are quite similar to the pooled estimates, although somewhat higher for mathematics than English or science.

### c. Human capital effect

To estimate the human capital effect, we compare the same teacher with him or herself over time as the teacher moves from future applicant to current applicant to past applicant. The model includes NB status indicators for whether the teacher is currently in, or has in the past participated in, the NB application process, along with teacher fixed effects, and student characteristics. Additional specifications include average baseline student test scores at the teacher level and either school characteristics or school fixed effects. The omitted NB status category is "future applicant," so the coefficient should be interpreted as the change in outcome score (in standard deviations from the national mean) resulting from having a teacher who is a current (or

past) NB applicant relative to having the same teacher at a stage in her or his career when she or he had not yet applied for certification. The coefficients should therefore pick up any effect on test scores from teachers who have gone through (past applicant), or are going through (current applicant), the NB certification process. The results of all subject areas are pooled due to the small number of teachers who change status in the certification process during the timeframe of analysis.

We present these results in Table 5 with the same layout as in the prior results tables. The coefficient estimates can be interpreted as the effect on test scores of having one semester with a teacher who is a current or past NB applicant relative to having that same teacher for a semester before she applies for NB certification. We find no evidence of a human capital effect; students of past or current applicants do not perform differently from students of the same teachers before they had applied for NB certification (future applicants). The effect sizes on both the current and past applicant indicator variables are small and not statistically different from zero in Kentucky or CPS for any specification.

## *VII.* ***Discussion and Conclusion***

Prior studies estimating the signaling and screening effects of NB use a variety of contexts and estimation strategies. They find modest relationships (on the order of 5 to 10 percent of a standard deviation) between student test scores and having a NB teacher. Corresponding estimates from this study tend to be on the lower end of that range. We also find strong evidence that controlling for school fixed effects or school-level characteristics is important, confirming that NB teachers are not randomly distributed across schools. Our estimates of the relationships between teacher NB status and student test scores are higher for students having NB teachers in $10^{th}$ and/or $11^{th}$ grade than for students having NB teachers in $8^{th}$ and/or $9^{th}$ grade. One conclusion might be that NB teachers assigned to higher level grades are better teachers than those assigned to teach

9th grade. Alternatively, these results are consistent with schools having more information about older students with which to sort them into classes, specifically information that is unobservable to the researcher. Either conclusion highlights the challenges of estimating impacts of NB teachers when students and teachers are not randomly assigned. To our knowledge, no previous study has been able to use longitudinal data to examine the human capital effects of participating in the NB certification process. Our ability to estimate such a model using teacher fixed effects mitigates against some of the concerns we note regarding nonrandom sorting of students and teachers. Indeed, our estimates suggest that there are no effects of the NB application process on teachers' ability to improve student test scores.

Significant financial nonfinancial resources go into the NB certification process. As of September 2005, the National Science Foundation and the U.S. Department of Education had appropriated more than $149 million dollars to it, and nongovernment funders had spent an additional $261 million (Cohen & Rice, 2005). Applicants for certification (and typically, their sponsoring school systems) also incur substantial costs. Many districts subsidize application fees and pay bonuses to successful candidates. For example, Goldhaber, Perry, and Anthony (2004) report that for each applicant certified in 2000, North Carolina paid $2,300 per applicant in assessment fees and $4,265 per certified teacher in terms of an annual salary increase. For applicant teachers, costs include time in addition to any direct financial costs of applying. District staff in CPS estimated that applicants could expect to spend 300-400 hours completing a NB application.

In spite of the financial and nonfinancial resources by teachers going through the application process and by districts and states encouraging and assisting teachers to become certified, we find no evidence that participating in the process makes teachers more effective at improving student test scores. NB certification was not designed as a professional development

program and might not be expected to change teacher practice. Given the time and resources devoted to completing the application process and the financial and nonfinancial supports provided by districts and states to promote NB certification, we would hope that the process might indeed improve teacher practice. However, future studies might consider a broader set of outcomes and might find evidence that the NB application process indeed changes teacher practice.

Table 1. Student characteristics by NB status

| | Chicago Public Schools | | | Kentucky | | |
|---|---|---|---|---|---|---|
| | Had an ever-achieved NB teacher | Had an unsuccessful NB applicant | Had only non-applicant teachers | Had an ever-achieved NB teacher | Had an unsuccessful NB applicant | Had only non-applicant teachers |
| | 9th Grade Students | | | 8th/9th Grade Students | | |
| Proportion Black | 0.36 | 0.46 | 0.39 | 0.15 | 0.13 | 0.09 |
| Proportion Latino | 0.38 | 0.41 | 0.48 | 0.03 | 0.04 | 0.02 |
| Proportion White | 0.16 | 0.08 | 0.08 | 0.79 | 0.80 | 0.86 |
| Proportion Male | 0.45 | 0.46 | 0.48 | 0.50 | 0.52 | 0.50 |
| Proportion FRPL[a] | 0.71 | 0.81 | 0.84 | 0.41 | 0.46 | 0.51 |
| Proportion classified as special education | 0.05 | 0.11 | 0.13 | 0.05 | 0.06 | 0.08 |
| Incoming test score (standardized) | 0.11 | -0.42 | -0.51 | 0.60 | 0.07 | -0.19 |
| Number of students | 15,592 | 7,733 | 46,416 | 2,858 | 1,546 | 76,086 |
| | 10th/11th Grade Students | | | 10th/11th Grade Students | | |
| Proportion Black | 0.32 | 0.57 | 0.45 | 0.14 | 0.17 | 0.08 |
| Proportion Latino | 0.45 | 0.34 | 0.44 | 0.03 | 0.03 | 0.02 |
| Proportion White | 0.14 | 0.05 | 0.07 | 0.76 | 0.73 | 0.85 |
| Proportion Male | 0.44 | 0.46 | 0.48 | 0.49 | 0.48 | 0.50 |
| Proportion FRPL[a] | 0.52 | 0.63 | 0.62 | 0.35 | 0.44 | 0.45 |
| Proportion classified as special education | 0.05 | 0.10 | 0.16 | 0.04 | 0.06 | 0.07 |
| Incoming test score (standardized) | 0.04 | -0.47 | -0.51 | 0.61 | 0.02 | -0.12 |
| Number of students | 20,482 | 8,102 | 19,962 | 6,939 | 2,512 | 105,014 |

Notes. The categories of students represented by the columns are mutually exclusive. If a student had both a successful teacher and an unsuccessful teacher, that student is included in the first column. Teachers who are first-time applicants during the last year of our analytic data (2011-12) are coded as their eventual outcome (i.e., successful or not successful) for the purposes of the descriptive tables. Teachers who apply but always withdraw are included as non-successful applicants.

[a] Free or reduced-price school lunch.

Table 2. Chicago Public Schools (CPS) teacher characteristics by NB status

| | Teacher is ever a successful applicant | Teacher is never a successful applicant | Teacher is never an applicant |
|---|---|---|---|
| | 9th Grade Teachers | | |
| Proportion Black | 0.13 | 0.33 | 0.23 |
| Proportion Latino | 0.07 | 0.10 | 0.08 |
| Proportion White | 0.70 | 0.49 | 0.47 |
| Proportion Male | 0.31 | 0.37 | 0.31 |
| Proportion Math or Science Major | 0.45 | 0.41 | 0.38 |
| Proportion Masters (highest degree) | 0.80 | 0.72 | 0.64 |
| Proportion Doctorate (highest degree) | 0.03 | 0.05 | 0.02 |
| Years of experience (in CPS) | 7.81 | 10.27 | 7.65 |
| Number of teachers | 215 | 181 | 2,518 |
| | 10th and 11th Grade Teachers | | |
| Proportion Black | 0.12 | 0.31 | 0.24 |
| Proportion Latino | 0.07 | 0.11 | 0.07 |
| Proportion White | 0.71 | 0.50 | 0.49 |
| Proportion Male | 0.30 | 0.41 | 0.32 |
| Proportion Math or Science Major | 0.41 | 0.45 | 0.37 |
| Proportion Masters (highest degree) | 0.80 | 0.71 | 0.64 |
| Proportion Doctorate (highest degree) | 0.03 | 0.06 | 0.02 |
| Years of experience (in CPS) | 7.81 | 10.22 | 8.32 |
| Number of teachers | 259 | 200 | 3,002 |

Notes. Teachers who are first-time applicants during the last year of our analytic data (2011-12) are coded as their eventual outcome (i.e., successful or not successful) for the purposes of the descriptive tables. Teachers who apply but always withdraw are included as non-successful applicants.

Table 3. Signaling Effect of NB Certification

| | Kentucky Effect of having at least one NB teacher in Grades 8/9 on test scores | | | Chicago Public Schools Effect of having at least one NB teacher in Grade 9 on test scores | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1') | (2') | (3') |
| Pooled (all subjects combined) | 0.042** | 0.013 | 0.010 | 0.153*** | 0.016 | 0.002 |
| | (0.018) | (0.013) | (0.366) | (0.033) | (0.018) | (0.028) |
| English | -0.004 | -0.002 | 0.000 | 0.055** | 0.016 | 0.036 |
| | (0.025) | (0.020) | (0.017) | (0.027) | (0.024) | (0.022) |
| Math | 0.122*** | 0.056** | 0.070*** | 0.132*** | 0.014 | -0.018 |
| | (0.034) | (0.024) | (0.018) | (0.047) | (0.027) | (0.024) |
| Science | 0.032 | 0.005 | -0.015 | 0.305*** | 0.023 | -0.034 |
| | (0.028) | (0.027) | (0.026) | (0.075) | (0.029) | (0.036) |
| | Kentucky Effect of having at least one NB teacher in Grades 10/11 on test scores | | | Chicago Public Schools Effect of having at least one NB teacher in Grades 10/11 on test scores | | |
| | (1) | (2) | (3) | (1') | (2') | (3') |
| Pooled (all subjects combined) | 0.071*** | 0.034** | 0.038*** | 0.159*** | 0.050*** | 0.048*** |
| | (0.022) | (0.015) | (0.012) | (0.017) | (0.012) | (0.011) |
| English | 0.076** | 0.028 | 0.026* | 0.105*** | 0.033** | 0.048*** |
| | (0.032) | (0.019) | (0.016) | (0.020) | (0.013) | (0.014) |
| Math | 0.099*** | 0.056** | 0.078*** | 0.200*** | 0.092*** | 0.069*** |
| | (0.038) | (0.024) | (0.009) | (0.027) | (0.024) | (0.022) |
| Science | 0.040 | 0.006 | 0.026 | 0.192*** | 0.012 | 0.012 |
| | (0.038) | (0.038) | (0.030) | (0.033) | (0.023) | (0.021) |
| Additional controls: | | | | | | |
| Student characteristics | X | X | X | X | X | X |
| Teacher experience (or proxy) | X | X | X | X | X | X |
| School characteristics | | X | | | X | |
| District characteristics | | X | | | n/a | |
| School fixed effects | | | X | | | X |
| Average incoming test score | | X | X | | X | X |

Notes. Sample size for grade 9 in CPS is 209,223 student-subject area observations with 69,741 students in each subject and covers 99 schools and 2,914 teachers. Sample size for grade 10/11 in CPS is 145,638 student-subject area observations with 48,546 students in each subject and covers 118 schools and 3,461 teachers. Sample size for grades 8/9 in Kentucky is 240,679 student-subject area observations with 80,263 students in English, 80,253 in math, and 80,163 in science. The grade 8/9 Kentucky sample covers 338 schools and 3,725 teachers. Sample size for grade 10/11 in Kentucky is 341,946 student-subject area observations with 114,019 students in English, 114,004 in math, and 113,923 in science. The grade 10/11 Kentucky sample covers 313 schools and 5,312 teachers. Student characteristics are race/ethnicity, gender, an indicator for FRPL eligibility, and an indicator for special education status. We control for teacher experience and experience-squared. Our proxy for teacher years of experience is years of experience teaching in CPS or number of years observed in the Kentucky data. School characteristics are percent of students of different race/ethnicity categories, percent of students eligible for FRPL, percent special education students, the natural logarithm of total enrollment, and school-level average composite ACT scores. Average

incoming test score is the average baseline student test score for students of the current year teacher. Standard errors are clustered at current teacher level.

Table 4. The NB Application Process as a Screening Mechanism

| | Kentucky<br>Effect of having one semester with an ever-certified NB teacher in Grades 8/9 on test scores compared to a semester with an applicant who is never certified | | | Chicago Public Schools<br>Effect of having one semester with an ever-certified NB teacher in Grade 9 on test scores compared to a semester with an applicant who is never certified | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1') | (2') | (3') |
| Pooled (all subjects combined) | 0.024*<br>(0.013) | 0.004<br>(0.012) | 0.007<br>(0.011) | 0.071***<br>(0.017) | 0.007<br>(0.011) | 0.007<br>(0.010) |
| English | -0.025<br>(0.016) | -0.019<br>(0.014) | -0.014<br>(0.012) | 0.029<br>(0.018) | -0.002<br>(0.015) | 0.025**<br>(0.012) |
| Math | 0.085**<br>(0.029) | 0.039<br>(0.024) | 0.053*<br>(0.022) | 0.069***<br>(0.026) | 0.011<br>(0.017) | 0.009<br>(0.019) |
| Science | 0.037<br>(0.024) | 0.011<br>(0.023) | 0.006<br>(0.020) | 0.121***<br>(0.037) | 0.003<br>(0.016) | −0.015<br>(0.020) |
| | Kentucky<br>Effect of having one semester with an ever-certified NB teacher in Grades 10/11 on test scores compared to a semester with an applicant who is never certified | | | Chicago Public Schools<br>Effect of having one semester with an ever-certified NB teacher in Grades 10/11 on test scores compared to a semester with an applicant who is never certified | | |
| | (1) | (2) | (3) | (1') | (2') | (3') |
| Pooled (all subjects combined) | 0.024<br>(0.015) | 0.012<br>(0.012) | 0.020*<br>(0.010) | 0.047***<br>(0.011) | 0.023***<br>(0.007) | 0.022***<br>(0.007) |
| English | 0.022<br>(0.021) | 0.022<br>(0.021) | 0.022<br>(0.021) | 0.028<br>(0.019) | 0.010<br>(0.013) | 0.019*<br>(0.010) |
| Math | 0.022<br>(0.026) | 0.005<br>(0.014) | 0.036*<br>(0.014) | 0.081***<br>(0.015) | 0.038***<br>(0.015) | 0.035**<br>(0.015) |
| Science | 0.013<br>(0.026) | 0.003<br>(0.020) | 0.019<br>(0.017) | 0.045**<br>(0.017) | 0.009<br>(0.010) | 0.018<br>(0.011) |
| Additional controls: | | | | | | |
| Student characteristics | X | X | X | X | X | X |
| Teacher experience (or proxy) | X | X | X | X | X | X |
| School characteristics | | X | | | X | |
| District characteristics | | X | | | | |
| School fixed effects | | | X | | | X |
| Average incoming test score | | X | X | | X | X |

Notes. See notes for Table 3.

Table 5. Human Capital Effects from Participating in the NB Application Process

| | | | | Chicago Public Schools Effect of the application process on student test performance (Grade 9) | | |
|---|---|---|---|---|---|---|
| | | | | (1') | (2') | (3') |
| *Pooled (all subjects combined)* | | | | | | |
| Current Applicant (Grade 9) | | | | 0.008 | 0.011 | 0.010 |
| | | | | (0.012) | (0.012) | (0.011) |
| Past Applicant (Grade 9) | | | | 0.002 | -0.009 | -0.005 |
| | | | | (0.015) | (0.015) | (0.014) |
| | Kentucky Effect of the application process on student test performance (Grades 10/11) | | | Chicago Public Schools Effect of the application process on student test performance (Grades 10/11) | | |
| | (1) | (2) | (3) | (1') | (2') | (3') |
| *Pooled (all subjects combined)* | | | | | | |
| Current Applicant (Grade 10) | -0.265 | -0.277 | -0.271 | 0.003 | -0.004 | 0.003 |
| | (0.261) | (0.261) | (0.265) | (0.012) | (0.012) | (0.012) |
| Past Applicant (Grade 10) | -0.022 | -0.027 | -0.034 | -0.014 | -0.022 | -0.015 |
| | (0.053) | (0.055) | (0.056) | (0.014) | (0.015) | (0.014) |
| Current Applicant (Grade 11) | 0.061 | 0.019 | 0.042 | 0.006 | -0.002 | 0.005 |
| | (0.074) | (0.056) | (0.080) | (0.014) | (0.015) | (0.014) |
| Past Applicant (Grade 11) | -0.003 | -0.018 | -0.018 | 0.003 | -0.013 | -0.005 |
| | (0.038) | (0.046) | (0.055) | (0.017) | (0.017) | (0.017) |
| Additional controls: | | | | | | |
| Student characteristics | X | X | X | X | X | X |
| School characteristics | | X | | | X | |
| District characteristics | | X | | | | |
| School fixed effects | | | X | | | X |
| Average incoming test score | | X | X | | X | X |

Notes. See notes for Table 3.

**References**


Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95–135.

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 No. 14608). Cambridge, MA: National Bureau of Economic Research.

Carnegie Task Force on Teaching as a Profession. (1986). *A nation prepared: Teachers for the 21st century.* New York: Carnegie Corporation of New York, Forum on Education and the Economy.

Cavalluzzo, L. (2004). *Is National Board certification an effective signal of teacher quality?* Alexandria, VA: The CNA Corporation.

Chicago Public Schools (2014) *CPS Teachers Achieve National Board Certification* Accessed at http://www.cps.edu/spotlight/pages/Spotlight548.aspx on October 16, 2014.

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.

Cohen, C., & Rice, J. K. (2005). *National Board certification as professional development: Pathways to success.* Washington, DC: The Finance Project.

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.

Elliott, S. W., Koenig, J. A., & Hakel, M. D. (2008). Assessing accomplished teaching: Advanced-level certification programs. Washington, DC: National Academies Press.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

Goldhaber, D. (2002). The mystery of good teaching. *Education Next, 2*(1), 50–55.

Goldhaber, D., Perry, D., and Anthony, E. (2004). The National Board for Professional Teaching Standards (NBPTS) Process: Who Applies and What Factors Are Associated with NBPTS Certification? *Educational Evaluation and Policy Analysis*, 26(4), 259-280.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, *42*(4), 765–794.

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

Harris, D. N., & Sass, T. R. (2006). *The effects of teacher training on teacher value-added.* Florida State University. Unpublished manuscript.

Harris, D. N., & Sass, T. R. (2007). *Teacher training, teacher quality, and student achievement.* Washington, DC: The Urban Institute, National Center for Analysis of Longitudinal Data in Education Research (CALDER).

Horoi, I., & Bhai, M. (2018). New evidence on national board certification as a signal of teacher quality. *Economic Inquiry*, *56*(2), 1185-1201.

Jacob, B. A., & Lefgren, L. (2004). The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in Chicago. *The Journal of Human Resources*, *50*(1), 50–79.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27,* 615–631.

Kastberg, David, Jessica Ying Chan, and Gordan Murray (2016). "Performance of U.S. 15-Year-Old Students in Science, Reading, and Mathematics Literacy in an International Context: A First Look at PISA 2015." U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved September 1, 2019 from http://nces.ed.gov/pubsearch.

McCaffrey, D. F., & Rivkin, S. (2007). *Empirical investigations of the effects of National Board of Professional Teaching Standards certified teachers on student outcomes.* Unpublished manuscript.

National Board for Professional Teaching Standards. (2019). *Guide to National Board Certification.* Arlington, VA Retrieved from http://www.nbpts.org/wp-content/uploads/Guide_to_NB_Certification.pdf

National Board Resource Center (2012). *Illinois NBCT Data December 2011*. Downloaded from http://nbrc.illinoisstate.edu/downloads/nbrc/NBCTReportData2012_000.pdf on October 16, 2014.

National Center for Education Statistics (2015). *2015: Mathematics and Reading at Grade 12*. Institute of Education Sciences, U.S. Department of Education. Retrieved from https://www.nationsreportcard.gov/reading_math_g12_2015/#1, 2019.

Podgursky, M., Springer, M. G., & Hutton, R.(2010). Teacher training and preparation in the United States. In D. J. Brewer & P. J. McEwan (Eds.), *Economics of education* (pp. 290–295). Amsterdam: Elsevier.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica, 73*(2), 417–458.

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

Sanders, W. L., Ashton, J. J., & Wright, S. P. (2005). *Comparison of the effects of NBPTS certified teachers with other teachers on the rate of student academic progress. Final report.* Arlington, VA: National Board for Professional Teaching Standards.

Smylie, M. A., Allensworth, E., Greenberg, R. C., Harris, R., & Luppescu, S. (2001). *Teacher professional development in Chicago: Supporting effective practice.* Chicago, IL: Consortium on Chicago School Research.

Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement.* Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.