THE IMPACT OF NATIONAL BOARD CERTIFIED TEACHERS ON STUDENT ACHIEVEMENT IN FLORIDA AND NORTH CAROLINA

A Summary of the Evidence Prepared for the National Academies Committee on the Evaluation of the Impact of Teacher Certification by NBPTS

Helen F. Ladd
Terry Stanford Institute
Duke University

Tim R. Sass
Dept. of Economics
Florida State University

Douglas N. Harris
Dept. of Ed. Policy Studies
University of Wisconsin

February 28, 2007

I. Introduction

Given the growing evidence that teacher quality plays a central role in determining student achievement (Rivkin, Hanushek and Kain (2005) and Rockoff (2004)), many states are looking for new ways to identify and reward effective teachers. Until recently, the primary state-level mechanism to improve teacher quality has been teacher certification systems that use academic degree requirements and written tests to limit entry into teaching. While every state has adopted such a system, evidence suggests certification does little to ensure that only effective teachers end up in classrooms.[1] As a result, policymakers are looking to alternatives, such as paying teachers based on specific types of knowledge or student outcomes such as student test score gains and related outcome measures.[2]

Another new approach to improving teacher quality comes from the National Board for Professional Teaching Standards (NBPTS) and represents a combination of certification and pay-for-performance. Unlike state systems, NBPTS certification is strictly voluntary, relatively difficult to obtain, and often encouraged by large financial incentives for those teachers who obtain it. Those teachers who apply go through an extensive process, taking paper-and-pencil tests and preparing and submitting lesson plans and videos of classroom instruction. These materials are reviewed and evaluated by the National Board which then determines which teachers will be certified. Those who receive the certification—only about half of the applicants—receive increases in salary through local and state policies that seek to encourage teachers to take part.

---

[1] See Hanushek (1986, 1997), Harris and Sass (2006a) and Angrist and Guryan (2004).

[2] At least one statewide merit pay plan has been tried (Tennessee) and another is being implemented (Florida). More common are district-level merit pay systems (e.g., Dallas, Denver, Houston). While most of these systems are too new to evaluate their efficacy, Dee and Keys (2004) provide an analysis of the impact of Tennessee's system on student achievement.

To promote teacher quality, NBPTS certification must accurately identify and certify teachers who are more effective than other teachers. If there is such a "signaling effect," then the certification and additional compensation provide incentives for teachers to work hard to become better teachers so that they can be certified and may also help to attract higher quality candidates into the profession. It is also possible that the process of NBPTS certification, which involves hundreds of hours of work, directly improves teacher effectiveness. This suggests that, in addition to the signaling effect, there may be a "human capital effect" of NBPTS certification.

In this report we provide evidence on both the signaling and human-capital effects of NBPTS certification, using data for both Florida and North Carolina. These are the only two states where individual students and teachers can be matched to specific classrooms statewide. In the following two sections we briefly describe the empirical models and data. We then present our findings., with particular attention to the differences in results that emerge from different specification of the model and those that emerge because of underlying differences between the two states. We conclude with a summary of the evidence and the implications for policy.

II. Model Specification

In order to gauge the impact of NBPTS certification on student achievement we begin with a general specification of the "value-added" model that relates gains in student achievement to vectors of time-varying student/family inputs (X), classroom-level inputs (C), school inputs (S) and time-invariant student/family characteristics (ψ):

$$A_{it} - A_{it-1} = \Delta A_{it} = \boldsymbol{\rho}_1 \mathbf{X}_{it} + \boldsymbol{\rho}_2 \mathbf{C}_{ijmt} + \boldsymbol{\rho}_3 \mathbf{S}_{mt} + \psi_i + \varepsilon_{it} \quad (1)$$

The subscripts denote individuals (i), classrooms (j), schools (m) and time (t).

Equation (1) is a restricted form of the cumulative achievement function specified by Todd and Wolpin (2003) where the achievement level at time t depends on the individual's initial endowment (eg. innate ability) and their entire history of individual, family and schooling inputs. Although often not stated, there are a number of implicit assumptions underlying the value-added model function specified in (1). First, it is assumed that the cumulative achievement function does not vary with age, is additively separable and linear. Second, family inputs are constant over time and the impact of these parental inputs on achievement, along with the effect of the initial individual endowment on achievement, change at constant rates. This allows the combination of these time-invariant inputs to be represented by the student-specific fixed component, ψ_i. Third, the marginal impacts of all prior school inputs decline geometrically with the time between the application of the input and the measurement of achievement at the same rate. Thus lagged achievement serves as a sufficient statistic for all prior schooling inputs. Fourth, school inputs each have an immediate one-time impact on achievement that does not decay over time.[3] A thorough discussion of these assumptions and the derivation of the linear value-added model can be found in Todd and Wolpin (2003) and Harris and Sass (2006b).

The vector of classroom inputs can be divided into four components: peer characteristics, $\mathbf{P}_{-ijmt}$ (where the subscript –i students other than individual i in the classroom); time-varying teacher characteristics (eg. experience and certification), $\mathbf{T}_{kt}$ (where k indexes teachers); time-invariant teacher characteristics (eg. innate ability and pre-service education), δ_k; and non-

[3] Thus, for example, the quality of a child's kindergarten must have the same impact on their achievement at the end of age 5 as it does on their achievement at age 18. While a strong assumption, this allows the impact of all prior schooling inputs to be captured by the lagged achievement score, A_{it-1}, on the left-hand side of the equation. Otherwise, equation (1) would contain a lagged dependent variable on the right hand side and thus could not be consistently estimated by ordinary least squares. In other work, Harris and Sass (2006b), we find that the degree of assumed persistence in the effect of prior schooling inputs has little effect on estimates of teacher effectiveness.

teacher classroom-level inputs (such as books, computers, etc.), $\mathbf{Z}_j$. If we assume that, except for teacher quality, there is no variation in education inputs across classrooms within a school, the effect of $\mathbf{Z}$ becomes part of the school-level input vector, $\mathbf{S}_m$.

The value-added model can then be expressed as:

$$A_{it} - A_{it-1} = \Delta A_{it} = \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{T}_{kt} + \delta_k + \boldsymbol{\beta}_4 \mathbf{S}_{mt} + \psi_i + \nu_{it} \quad (2)$$

where ν_{it} is a normally distributed, mean zero error.

Of particular interest for this study is the effect on student achievement of being taught by a board certified teacher. Thus, in our simplest specification, we distinguish teachers according to whether or not they ever receive NBPTS certification.[4] We denote the average fixed effect for teachers who never become NBPTS certified as $\bar{\delta}^{\text{never NBCT}}$, which is simply a constant, α. We then denote the difference in the average fixed effect for teachers who at some point become NBCTs and those who never become NBCTs $(\bar{\delta}^{\text{ever NBCT}} - \bar{\delta}^{\text{never NBCT}})$ as γ(Ever NBCT), where (Ever NBCT) is an indicator variable representing teachers who become NBCTs at some point in time. The average fixed effect for teachers who become NBCTs at some point is thus $\alpha+\gamma$(Ever NBCT). Our value-added model can re-written as:

[4] By comparing ever-NBCTs with never-NBCTs, we are comparing teachers who apply for and receive NBPTS certification with those who either never apply or who apply but fail to be certified. This comparison is relevant for determining if the voluntary system currently in place rewards teachers who are more effective in boosting student achievement. Alternatively, one could compare unsuccessful applicants and successful applicants. This comparison would indicate if the certification process is successful in sorting out superior teachers within the group of applicants. This would be relevant if a state mandated that all teachers apply for NBPTS certification. However, we are unaware of any states that are considering making the NBPTS certification process mandatory.

$$\Delta A_{it} = \alpha + \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{T}_{kt} + \gamma(Ever\ NBCT) + \boldsymbol{\beta}_4 \mathbf{S}_{mt} + \psi_i + \nu_{it} \quad (3)$$

If the average effectiveness of future NBCTs exceeds the average of teachers who never become NBCTs then γ would be positive.

The effectiveness of teachers who become NBPTS certified could change during the certification process. Following the work of Goldhaber and Anthony (forthcoming), it may be that teacher productivity temporarily falls during the certification process, but later recovers. In addition, it may be that the certification process itself enhances future teacher productivity. If the certification process leads teachers to re-evaluate their teaching methods or if preparation for the exam components causes teachers to sharpen their content knowledge then their effectiveness could rise after the certification process. To account for these possible intertemporal changes in the effectiveness of teachers who become NBPTS-certified we interact the (Ever NBCT) variable with indicator variables denoting the three time periods: pre-application, application year and received certification. This extension yields:

$$\begin{aligned}\Delta A_{it} = \alpha + \boldsymbol{\beta}_1 \mathbf{X}_{it} + \boldsymbol{\beta}_2 \mathbf{P}_{-ijmt} + \boldsymbol{\beta}_3 \mathbf{T}_{kt} + \gamma_1(Ever\ NBCT \times \text{Pre-Application}) + \\ \gamma_2(Ever\ NBCT \times \text{Application Year}) + \gamma_3(Ever\ NBCT \times \text{ReceivedCertification}) + \boldsymbol{\beta}_4 \mathbf{S}_{mt} + \\ \psi_i + \nu_{it}\end{aligned} \quad (4)$$

The coefficients γ_1, γ_2 and γ_3 represent the difference between the average effectiveness of teachers who are ever-NBPTS-certified and those who are never-NBPTS-certified during the relevant periods. Tests of whether these are statistically different from zero provide evidence regarding the signaling hypothesis. The human capital hypothesis can be tested by comparing the differences in coefficients, ie. γ_2-γ_1 and γ_3-γ_1.

Equation (4) is our "baseline" model, which we employ in most of the analyses. The vector of time-varying student characteristics, $\mathbf{X}_{it}$, includes indicators for repeating a grade, a "structural" move from another school and a "non-structural" move from another school. Classroom peer characteristics, $\mathbf{P}_{-ijmt}$, include class size and the proportion of classmates who are non-white. Teacher characteristics, $\mathbf{T}_{kt}$, include both time-varying and time-invariant teacher attributes. The time-varying teacher characteristics are: six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher time-invariant demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. No explicit school-level variables are included in the estimation of the baseline version of equation (4). School fixed effects are omitted in the baseline because comparing NBPTS teachers only to other teachers within their school may give a misleading picture of the overall effectiveness of these teachers, depending on how teachers are sorted across schools. Complete variable definitions are provided in the appendix.

In addition to the "baseline" specification we estimate numerous alternative specifications as robustness checks. These alternative specifications include:

- Model "B" - use of student covariates rather than student fixed effects to control for student heterogeneity
- Model "C" - exclusion of state teacher certification and advanced degree indicators
- Model "D" - use of student covariates rather than student fixed effects to control for student heterogeneity; exclusion of state teacher certification and advanced degree indicators
- Model "E" - inclusion of school fixed effects to control for unmeasured school quality

- Model "F" - use of student covariates rather than student fixed effects to control for student heterogeneity; inclusion of school fixed effects to control for unmeasured time-invariant school quality
- Model "G" - inclusion of school fixed effects to control for unmeasured school quality; exclusion of state teacher certification and advanced degree indicators
- Model "H" - use of student covariates rather than student fixed effects to control for student heterogeneity; inclusion of school fixed effects to control for unmeasured school quality; exclusion of state teacher certification and advanced degree indicators

For each of the alternative specifications that replace student fixed effects with student time-invariant controls the vector of student demographics includes: race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Regressions for students in the elementary grades also include the student's age the first time in grade 3.

It is important to emphasize that all comparisons between Florida and North Carolina are based on the common model specifications described above. Thus any observed differences in the effectiveness of NBPTS-certified teachers (NBCTs) across states are not attributable to differences in the empirical specifications employed.

III. Data and Sample Selection

The data requirements necessary to estimate the above models are demanding. Not only must individual students be tracked over time, but students and teachers must also be matched to specific classrooms. The necessary data are currently available on a statewide basis only in Florida and North Carolina. In Florida, students and teachers can be matched at all grade levels

whereas in North Carolina they can be reliably matched only at the elementary school level.. Thus we conduct all of the cross-state comparisons at the elementary school level. We supplement those comparisons with additional analyses at the middle and high school level for Florida..

The primary source of data covering Florida teachers and students is the Florida Department of Education's K-20 Education Data Warehouse (EDW), an integrated longitudinal database covering all Florida public school students and school employees from pre-school through college. In North Carolina, the data are derived from administrative records maintained by the North Carolina Education Research Data Center, housed at Duke University.

The states differ somewhat in their available achievement data. Florida has both a criterion-referenced standards-based exam and a norm-referenced test in grades 3-10. In North Carolina, only the curriculum-referenced "End-of-Grade" exam is available and only fpr grades 3-8. To align the empirical estimates as closely as possible, the Florida analysis is limited to the criterion-refenced exam.[5] To avoid problems with differing achievement scales in each state we normalize both sets of achievement scores to standard deviation units by grade by year.

In North Carolina statewide achievement testing records for elementary grades are available from 1994/95 through 2003/04. In contrast, statewide criterion-referenced tests in Florida were first administered in the 2000/01 school year. To make the analysis comparable, we limit our cross-state analyses to achievement scores in the 2000/01 through 2003/04 school years. This period translates to three years of achievement gains, 2001/02-2003/04. We also

[5] These two tests are also the basis for a variety of accountability measures, intended to provide incentives for students and educators to improve performance. For this reason, it is reasonable to expect that the exam aligns well with the curriculum taught in schools and is therefore sensitive to changes in the quality of instruction.

include some additional analyses of North Carolina that cover a longer time period of achievement gains , 1995/96 – 2003/04.

We place two additional restrictions on the sample we analyze. First, to avoid atypical classroom settings we consider only courses in which 10-50 students are enrolled. Second, we eliminate students enrolled in charter schools from the analysis since they may have differing curricular emphases and student-peer and student-teacher interactions may differ in fundamental ways from traditional public schools.

Information on NBPTS certification is available through calendar year 2004. Thus we can determine if individuals teaching in 2003/04 later achieved certification in the Fall of 2004. Tables 1A-C provide summary statistics on newly certified NBCTs by year of certification. In Florida the proportion of teacher receiving certification who were black increased over time from 0.0 percent in 1999 to about 5-1/2 percent in 2004. In North Carolina, the proportion of newly-minted NBCTs who were black was relatively constant between 1999 and 2002 then dropped considerably in 2003 and 2004. In Florida, early cohorts of newly-certified NBCTs tended to be somewhat more experienced than later cohorts and a much larger proportion possessed advanced degrees (7 percent in 1999 versus 3.6 percent in 2004). A similar pattern is observed in North Carolina with a general decline in experience and possession of advanced degrees for new NBCTs from 1999 through 2004.

IV. Results

We first report estimates of the basic model, and the various alternative specifications of that model, using a common time period for the two states. We then present some variations of

the model that because of data limitations, can only be estimated for one or the other of the two states. .

Table 2 shows estimates of the basic model with a single indicator variable for whether a teacher is ever NBPTS certified during the years under study or in previous years (equation (3)). The first thing to notice is the striking similarity in the estimated coefficients of the non-NBCT characteristics of teachers across the two states. For example, the estimated impact of experience is quite similar in both states, though the marginal effect of experience past the first couple of years tends to be somewhat higher in North Carolina than in Florida. In addition, emerging for both states is a somewhat surprising negative relationship between possession of advanced degree and gains in student achievement. Finally, in both states black teachers have lower effects on student achievement gains in math than the average teacher whereas, if anything, black teachers tend to outperform the average teacher in promoting reading achievement in both Florida and North Carolina.

The results from this baseline model suggest that ever-NBPTS teachers are more effective than other teachers in North Carolina, boosting test scores by 4 percent of a standard deviation in reading and 7 percent of a standard deviation in math. In Florida, the results are statistically significant only in reading and, even in this case, the point estimate is considerably smaller than in North Carolina (about 2 percent of a standard deviation). Because the model specifications and time period are identical, these differences in estimated effects of national board certitication are not attrirbutable to differences in those factors, a conclusion reinforced by the remaining results.

To separate the signal and human capital effects, we break the ever-NBPTS variable into three indicator variables indicating three time periods: two or more years before they become

certified, the application year (ie. the year prior to certification) and the period after NBPTS certification is obtained. The first represents an initial signaling effect, indicating whether the types of teachers who become NBPTS certified are inherently more effective. The *difference* between the impact of the second indicator and the first indicates the effect of preparing for certification on teacher effectiveness. Similarly, the *difference* in the impacts of the third and first indicators measures the net change in teacher effectiveness from completing the NBPTS certification process. Results from estimating this baseline model (equation (4)) are presented in Table 3.

We find a positive signaling effect for NBPTS teachers in North Carolina in both math and reading, though the magnitudes are not large. North Carolina teachers who later become NBCTs contribute 4 percent of a standard deviation more to student achievement in reading than the average North Carolina teacher who never becomes a NBCT during the sample period. For math the signaling effect in North Carolina is higher, about 6 percent of a standard deviation. For Florida we find a positive signaling effect in reading of about 5 percent of a standard deviation. However, we find no significant difference in the value added by Florida elementary math teachers who later become NBCTs compared to the average Florida elementary math teacher who is never observed to hold NBPTS certification.

There is no evidence that going through the process of NBPTS certification has a positive impact on teacher effectiveness. That is, we find no support for a "human capital effect." For Florida, there is a statistically significant decline in reading teacher effectiveness during the certification preparation period that remains post-certification. In North Carolina, the relevant coefficients for the application or post application year do not differ in a statistically significant way from the pre –certification coefficient.

Table 4 summarizes a variety of robustness checks.[6] Each cell indicates whether the results are qualitatively different when the indicated changes are made to the baseline model. "Yes" in any cell means that one of the three NBPTS coefficients changes its statistical significance (a significant coefficient becoming insignificant or vice versa). Full results for each of the alternative models can be found in the appendix. The only change that appears to have a meaningful impact is the addition of school fixed effects. Adding school fixed effects of course changes the comparison group from all teachers to other teachers within the same school. However, even in this case, the results are not dramatically different. As indicated in Table 5, the only change for North Carolina is that the application-year coefficient for reading achievement goes from being marginally significant to being insignificant. Similarly, the post-certification coefficient for math achievement in Florida goes from being statistically insignificant to marginally significant.

Table 6 shows results for model variations in which the persistence of past educational inputs is estimated from the data rather than assumed to have a coefficient of one.[7] Thus in these models the dependent variable is the contemporaneous achievement level and the lagged achievement score is included as a right-hand-side explanatory variable. Allowing partial persistence has little effect on the estimates for North Carolina. Only when school fixed effects are added is there any qualitative change at all, as discussed above. In contrast, there is considerable variation in the estimated impacts of NBCTs in Florida when the impact of prior inputs is allowed to decay over time. Table 7 shows that with partial persistence (and student fixed effects supplanted by student demographics) the post-certification effect of NBCTs in

[6] In addition to the plethora of model estimates reported in the text, the results from additional specifications, particularly ones based on aggregation to the grade-by-school-by-year level, can be found in the appendix.
[7] Complete results are provided in the appendix.

Florida is positive in both math and reading, though rather small (about 2 percent of a standard deviation). The pre-application signaling effect for Florida elementary reading teachers remains marginally significant, but the magnitude of the effect is cut in half.

As demonstrated by Todd and Wolpin (2003), the lagged dependent variable on the right hand side of the variable persistence model will be correlated with the error term, yielding biased ordinary least squares (OLS) coefficient estimates. To minimize this problem we parametrically assign a persistence value of 0.67 and re-estimate the models using OLS. As illustrated in Table 8, the results do not change significantly, suggesting that any bias in the variable persistence models is small.

Tables 9 and 10 present estimates of models corresponding to the variable-partial persistence and fixed partial persistence estimates presented in Tables 7 and 8. The difference is that the models whose estimates are presented in Tables 9 and 10 contain school fixed effects. The same basic patterns emerge. Allowing for variable persistence boosts the post-certification effects in Florida and makes little difference in the North Carolina results.

In addition to the cross-state comparisons, we can exploit features unique to the data in Florida or unique to the North Carolina data to further explore the robustness of our findings. In particular, since it is possible to match students and teachers to classrooms at all grade levels in Florida we can determine if the impacts of NBPTS certification differ between the elementary, middle and high-school levels. Because Florida administers two achievement tests, we can also see if our findings are sensitive to the type of test that is administered. Similarly, a longer time span of data available in North Carolina permits us to gauge the impact of changing the period of analysis on the results. Finally, available data on teachers' undergraduate institutions and state

certification exam scores allow us to see to what extent alternative measures can be used to detect differences in teacher quality.

Table 11 presents estimates of the baseline model (equation (4)) for the subsamples of elementary, middle and high-school students in Florida. For reading, the positive signaling effect found in elementary school also shows up in middle school, but not in high school. For math we find a positive signaling effect of 5 percent of a standard deviation in middle school, but no significant signaling effects in either elementary or high school. There is evidence of a small positive human-capital effect of NBPTS certification for high school math. However, the human-capital effect is negative for middle-school math and for middle-school reading. As with the elementary school results, the estimates for middle and high school are not particularly sensitive to the inclusion of school fixed effects (see Table 12).

Tables 13, 14 and 15 present estimates of the baseline model for elementary, middle and high school, respectively, using both the criterion-referenced Sunshine State Standards (SSS) exam and the norm-referenced Stanford-9 exam. At the elementary school level (Table 13) we find that the results are indeed sensitive to which exam is used to gauge student achievement and teacher value-added. Use of the Stanford-9 exam yields a smaller, though still statistically significant signaling effect for elementary school reading teachers who become NBCTs, relative to the SSS exam. In contrast, the SSS yields no statistically significant signaling effect in math whereas the Stanford-9 produces a positive and statistically significant signaling effect of about 4 percent of a standard deviation in elementary math. These results are comparable to those found for North Carolina (see Table 3). However, even when using the Stanford-9 exam the estimated human-capital effects are quite different across the two states. In North Carolina the estimated effectiveness of NBCTs is essentially the same before and after the certification process. In

contrast, the Florida estimates based on the Stanford-9 suggest a negative human-capital effect, wiping out any pre-application advantage of eventual NBCTs over never-NBCTs. In contrast to the elementary school results, the middle school estimates (Table 14) and the high school estimates (Table 15) are relatively robust to changes in the exam used to measure student performance.

Tables 16 and 17 present estimates for North Carolina math and reading achievement, allowing for the inclusion of additional measures of teacher quality and for a longer time period of analysis.[8] For math achievement (Table 16) we find that adding teachers' test scores and measures of the quality of their undergraduate institution to the model have virtually no impact on the estimated effects of NBPTS certification. Although the correlation between teacher exam scores and student achievement (controlling for NBPTS certification) is positive and statistically significant, the inclusion of teacher exam scores does not raise the overall fit of the model, as indicated by a constant R-squared value. None of the indicators of undergraduate institutional quality are statistically significant. Extending the time period of analysis also has little impact on the estimates of NBPTS certification effects for elementary math in North Carolina. For reading achievement (Table 17) the results are also quite robust to controls for teacher exam scores and undergraduate institutional quality. However, extending the time period of analysis from 99/00-03/04 to 96/97-03/04 does alter the precision of the signaling effect estimates. Whether or not additional teacher variables are included, the difference in value added between future NBCTs and the average teacher who is never observed to obtain NBPTS certification is statistically insignificant.

[8] The samples used in this analysis are somewhat different from those reported in Table 3, as can be readily seen from the differing number of observations. However, the resulting estimates are quite similar. Further, what is important for the robustness analysis is how the estimates vary across specifications for a given sample.

VII. Summary and Conclusions

In this report we use rich longitudinal databases from Florida and North Carolina to evaluate the efficacy of the program of voluntary national teacher certification established by NBPTS. There are two avenues by which NBPTS could promote teacher quality. First, NBPTS certification could serve as a tool to identify superior teachers and allocate rewards to them. We term this mechanism the "signaling effect." Alternatively, the NBPTS certification process could directly impact teacher quality by encouraging teachers to improve their skills or modify their teaching practices. This second avenue for improving teacher quality we call the "human capital effect."

In our primary analysis we estimated identical empirical models of student achievement in elementary school for both Florida and North Carolina, using the same time period of analysis. In this direct cross-state comparison we found strong evidence of a signaling effect of NBPTS certification in both math and reading for North Carolina. For Florida, a statistically significant signaling effect was only found for reading. For both states the magnitude of the signaling effects are not large, roughly 4-6 percent of a standard deviation in achievement. In neither state did we find strong evidence of a positive human-capital effect.

We tested the robustness of our primarly findings by estimating a multitude of alternative model specifications. For North Carolina the findings are extremely robust to changes in model specification. The Florida results are also quite robust, though they are sensitive to the assumptions that are made with respect to the persistence or prior schooling inputs.

Unique features of each state's data allowed us to also explore the effects of changing exams, grade levels, and time periods of analysis within a state. Using data from Florida we find that the positive signaling effect in elementary school also shows up in middle school, but not in

high school. For math we find a positive signaling effect in middle school, but no significant signaling effects in either elementary or high school. Since Florida administers two achievement tests in both math and reading each year we can determine how the estimated NBPTS effects vary with the test instrument in the Sunshine State. We find that the Florida results are indeed quite sensitive to the choice of test to gauge student achievement, particularly at the elementary school level. In North Carolina, data on teacher exam scores and measures of the quality of teachers' undergraduate institutions are available. Adding these alternative measures of teacher quality to the model had virtually no effect on the estimated impacts of NBPTS certification, however. The North Carolina data also cover a longer time span than the available data in Florida. Extending the period of analysis back in time did little to affect the estimated NBPTS certification effects in math but did reduce the precision of the estimates in reading, producing insignificant estimated signaling effects.

It seems clear from our analysis that in North Carolina, elementary teachers who eventually become NBCTs are more effective than the average teacher, prior to going through the certification process. The evidence on signaling effects in Florida is mixed, with results varying by grade level and by the exam used to gauge student progress. For both states, the signaling effects, when the exist, are of modest magnitude – about 4 to 6 percent of a standard deviation in student achievement. For both Florida and North Carolina, the preponderance of the evidence indicates that the process of NBPTS certification does not enhance teacher quality.

Why the ability of NBPTS certification to identify more effective teachers appears to vary between Florida and North Carolina is still not clear. Potentially the pool of teachers who choose to obtain certification could differ across states, though rudimentary analysis of the observable characteristics of NBCTs across the two states does not seem to support this.

Alternatively, the different achievement tests in the two states could be measuring different aspects of student learning and thus be measuring different margins of teacher effectiveness. What we do know, however, is that the observed differences are not attributable to differences in model specification or the time period of analysis.

References

Angrist, Joshua D. and Jonathan Guryan. 2004. "Teacher Testing, Teacher Education, and Teacher Characteristics." *American Economic Review* 94(2):241-246.

Dee, Thomas S., and Benjamin J. Keys. 2004. "Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment." *Journal of Policy Analysis and Management* 23(3):471-488.

Goldhaber, Dan, and Emily Anthony. Forthcoming. "Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching." *Review of Economics and Statistics*.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24(3): 1141-77.

Hanushek, Eric A. 1997. "Assessing the Effects of School Resources on Student Performance: An Update." *Educational Evaluation and Policy Analysis* 19(2):141-164.

Harris, Douglas, and Tim R. Sass. 2006a. The Effects of NBPTS Teachers on Teacher Value-Added. Unpublished, Florida State University.

Harris, Douglas, and Tim R. Sass. 2006b. The Effects of Teacher Training on Teacher Value-Added. Unpublished, Florida State University.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools and Academic Achievement." *Econometrica* 73(2):417-58.

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

Todd, Petra E. and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal* 113(485):F3-F33.

Table 1A
Florida – NBCT Summary Statistics by Certification Year – Math Elementary

	yr = 1999		yr = 2000		yr = 2001		yr = 2002		yr = 2003		yr = 2004	
Variable	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
Male	85	0.012	101	0.020	144	0.042	156	0.038	156	0.051	167	0.066
Black	85	0.000	101	0.000	144	0.056	156	0.032	156	0.038	167	0.054
Hispanic	85	0.047	101	0.079	144	0.125	156	0.077	156	0.090	167	0.090
Other Non-White	85	0.000	101	0.010	144	0.000	156	0.006	156	0.000	167	0.018
ever_nbpts_generalist	85	0.976	101	0.970	144	1.000	156	0.981	156	0.962	167	0.874
ever_nbpts math	85	0.000	101	0.000	144	0.000	156	0.000	156	0.000	167	0.000
ever_nbpts_special ed	85	0.000	101	0.020	144	0.000	156	0.000	156	0.006	167	0.006
nbpts_certification year	85	1999	101	2000	144	2001	156	2002	156	2003	167	2004
exp at time of NBPTS cert.	83	11.944	94	11.452	137	10.468	154	10.547	155	10.150	167	10.079
adv_degree	85	0.682	101	0.483	144	0.448	156	0.462	156	0.444	167	0.356

Table 1B
Florida – NBCT Summary Statistics by Certification Year – Reading Elementary

	yr = 1999		yr = 2000		yr = 2001		yr = 2002		yr = 2003		yr = 2004	
Variable	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
Male	92	0.011	110	0.027	161	0.043	170	0.047	168	0.054	190	0.074
Black	92	0.000	110	0.000	161	0.056	170	0.029	168	0.036	190	0.058
Hispanic	92	0.043	110	0.082	161	0.124	170	0.071	168	0.089	190	0.089
Other Non-White	92	0.000	110	0.018	161	0.000	170	0.006	168	0.000	190	0.016
ever_nbpts_generalist	92	0.978	110	0.973	161	1.000	170	0.982	168	0.964	190	0.868
ever_nbpts_langarts	92	0.022	110	0.009	161	0.000	170	0.000	168	0.006	190	0.005
ever_nbpts_special ed	92	0.000	110	0.018	161	0.000	170	0.000	168	0.006	190	0.011
nbpts_certification year	92	1999	110	2000	161	2001	170	2002	168	2003	190	2004
exp at time of NBPTS cert.	89	11.738	106	11.075	154	10.599	168	10.511	167	10.284	190	10.268
adv_degree	92	0.664	110	0.475	161	0.450	170	0.474	168	0.438	190	0.358

Table 1C
North Carolina – NBCT Summary Statistics by Certification Year – Elementary (Math and Reading)

Certification Year	# New NBCTs	Percent Male	Percent Black	Percent Hispanic	Percent Nonwhite	Mean Experience	% with Adv Degree	Mean Test Score
95	4	0.00%	0.00%	0.00%	0.00%	12.5	75.00%	0.084
96	8	12.50%	0.00%	0.00%	0.00%	17.1	75.00%	0.923
97	14	0.00%	0.00%	0.00%	0.00%	13.9	71.43%	0.524
98	42	0.00%	7.14%	0.00%	9.52%	11.8	64.29%	0.500
99	78	2.56%	7.69%	0.00%	7.69%	13.3	46.15%	0.271
2000	147	5.44%	5.44%	0.00%	6.12%	12.4	51.02%	0.462
2001	140	2.86%	8.57%	0.71%	10.00%	12.0	58.57%	0.258
2002	135	3.70%	8.89%	0.00%	8.89%	11.1	41.48%	0.363
2003	166	4.82%	4.82%	0.60%	5.42%	11.4	33.73%	0.263
2004	188	5.32%	4.26%	0.00%	4.26%	11.8	31.91%	0.286

Table 2
"Baseline Model" Estimates of the Effects of
NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina
(Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Ever a NBCT	0.0032	0.0665**	0.0153*	0.0384**
	(0.0082)	(0.0061)	(0.0073)	(0.0052)
1-2 Years of Experience	0.0555**	0.0469**	0.0658**	0.0206**
	(0.0088)	(0.0078)	(0.0081)	(0.0071)
3-5 Years of Experience	0.0521**	0.0837**	0.0497**	0.0692**
	(0.0093)	(0.0077)	(0.0084)	(0.0068)
6-12 Years of Experience	0.0622**	0.0902**	0.0543**	0.0754**
	(0.0090)	(0.0073)	(0.0081)	(0.0066)
13-20 Years of Experience	0.0722**	0.1086**	0.0661**	0.0883**
	(0.0093)	(0.0078)	(0.0084)	(0.0070)
21-27 Years of Experience	0.0579**	0.1104**	0.0763**	0.1002**
	(0.0099)	(0.0080)	(0.0089)	(0.0071)
28 or More Years of Experience	0.0488**	0.1053**	0.0716**	0.1071**
	(0.0106)	(0.0085)	(0.0097)	(0.0077)
Standard State License	0.0213*	0.0289**	0.0361**	-0.0028
	(0.0096)	(0.0069)	(0.0088)	(0.0062)
Advanced Degree	-0.0100*	-0.0122**	-0.0257**	-0.0179**
	(0.0043)	(0.0041)	(0.0039)	(0.0035)
Male Teacher	-0.0061	0.0026	-0.0177**	0.0043
	(0.0062)	(0.0064)	(0.0054)	(0.0055)
Black Teacher	-0.0156*	-0.0459**	0.0103+	0.0086+
	(0.0061)	(0.0058)	(0.0056)	(0.0052)
Hispanic Teacher	0.0342**	-0.0085	0.0812**	-0.0331
	(0.0087)	(0.0312)	(0.0082)	(0.0316)
Other Non-White Teacher	0.0014	-0.0199	0.0318	-0.0151
	(0.0199)	(0.0209)	(0.0215)	(0.0190)
Teacher Same Race as Student	0.0203**	0.0316**	0.0106*	0.0221**
	(0.0050)	(0.0045)	(0.0048)	(0.0047)
Teacher Same Gender as Student	-0.0009	-0.0010	0.0036	-0.0136**
	(0.0043)	(0.0040)	(0.0042)	(0.0045)

Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No
R-squared	0.611	0.587	0.668	0.574
No. of Student-Year Obs.	478599	597575	620886	602113

Note: dependent variable is the gain in the normalized (by grade and year) achievement test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina). For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus year 2000/01 gains for Florida include grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus year 2000/01 gains in elementary school for North Carolina include grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 3
"Baseline Model" Estimates of the Effects of
NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina
(Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Two or More Years Prior to Becoming an NBCT	-0.0069	0.0580**	0.0462**	0.0370**
	(0.0147)	(0.0145)	(0.0143)	(0.0116)
Application Year for NBCT	-0.0053	0.0442**	-0.0095	0.0235*
	(0.0169)	(0.0144)	(0.0156)	(0.0117)
Received Natl. Board Certification	0.0107	0.0737**	0.0091	0.0422**
	(0.0106)	(0.0071)	(0.0093)	(0.0061)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No
R-squared	0.611	0.587	0.668	0.574
No. of Student-Year Obs.	478599	597575	620886	602113

Note: dependent variable is the gain in the normalized (by grade and year) achievement test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina). For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus year 2000/01 gains for Florida include grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus year 2000/01 gains in elementary school for North Carolina include grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 4
Summary of Alternative Specifications to "Baseline Model"
Do Alternative Specifications Produce Qualitatively Different Estimates of the Effects of NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina?

Change in Model Specification (Relative to Baseline Model)	Florida		North Carolina	
	Math	Reading	Math	Reading
Remove Student FE, Add Student Demographics	No	No	No	No
Remove State Licensure and Advanced Degree Variables	No	No	No	No
Remove Student FE, Add Student Demo., Remove Licensure and Adv. Degree	No	No	No	No
Add School FE	***Yes***	No	No	***Yes***
Remove Student FE, Add Student Demo., Add School FE	No	***Yes***	No	No
Remove State Licensure and Advanced Degree Variables, Add School FE	***Yes***	No	No	***Yes***
Remove Student FE, Add Student Demo., Remove Licesure/Degree, add School FE	No	No	No	***Yes***

Note: a qualitative change is defined as a case where a coefficient that is statistically significant becomes insignificant, a coefficient that is insignificant becomes significant, or a coefficient is statistically significant in both specifications but changes sign. Comparisons are relative to "baseline" model.

Table 5
Estimates of the Effects of NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina – Baseline Model Plus School Fixed Effects (Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Two or More Years Prior to Becoming an NBCT	-0.0099 (0.0153)	0.0765** (0.0155)	0.0434** (0.0151)	0.0467** (0.0122)
Application Year for NBCT	0.0051 (0.0177)	0.0453** (0.0152)	-0.0086 (0.0165)	0.0152 (0.0125)
Received Natl. Board Certification	0.0217+ (0.0112)	0.0776** (0.0078)	0.0089 (0.0100)	0.0376** (0.0065)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.670	0.648	0.718	0.642
No. of Student-Year Obs.	478599	597575	620886	602113

Note: dependent variable is the gain in the normalized (by grade and year) achievement test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina). For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus year 2000/01 gains for Florida include grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus year 2000/01 gains in elementary school for North Carolina include grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 6
Summary of Effects of Allowing Partial Persistence in Past Schooling Inputs
Does Allowing Partial Persistence of Past Schooling Inputs
Produce Qualitatively Different Estimates of the Effects of NBPTS-Certified Teachers
On Student Achievement in Florida and North Carolina?
(Grades 4-5, 2000/01-2003/04)

Model Specification	Florida		North Carolina	
	Math	Reading	Math	Reading
Baseline Less Student FE, Plus Student Demographics	***Yes***	***Yes***	No	No
Baseline Less Student FE, Plus Student Demo., Less Licensure and Adv. Degree Variables	***Yes***	***Yes***	No	No
Baseline Less Student FE, Plus Student Demo., Plus School FE	***Yes***	***Yes***	No	***Yes***
Baseline Less Student FE, Plus Student Demo. Less Licensure & Adv. Deg., Plus School FE	***Yes***	***Yes***	No	***Yes***

Note: a qualitative change is defined as a case where a coefficient that is statistically significant becomes insignificant, a coefficient that is insignificant becomes significant, or a coefficient is statistically significant in both specifications but changes sign. Comparisons are relative to "baseline" model.

Table 7
Estimates of the Effects of NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina -- Baseline Model Less Student FE, Plus Student Demographics with Variable Persistence in Prior Schooling Inputs (Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Two or More Years Prior to Becoming an NBCT	0.0127 (0.0111)	0.0441** (0.0108)	0.0203* (0.0097)	0.0205* (0.0085)
Application Year for NBCT	0.0050 (0.0125)	0.0334** (0.0124)	0.0073 (0.0108)	0.0288** (0.0090)
Received Natl. Board Certification	0.0252** (0.0081)	0.0620** (0.0056)	0.0176** (0.0064)	0.0347** (0.0042)
Lagged Test Score	0.6794** (0.0016)	0.7298** (0.0012)	0.6742** (0.0015)	0.6894** (0.0012)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	Yes	Yes	Yes	Yes
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	No	No	No	No
School Fixed Effects	No	No	No	No
R-squared	0.646	0.721	0.635	0.665
No. of Student-Year Obs.	454054	580934	477382	585432

Note: dependent variable is the normalized (by grade and year) achievement level on the test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina). The lagged test score is included as a regressor. For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus in year 2000/01 lagged scores are available for grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus lagged scores are available in year 2000/01 in elementary school for North Carolina in grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 8
Estimates of the Effects of NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina -- Baseline Model Less Student FE, Plus Student Demographics with Persistence in Prior Schooling Inputs Fixed at 0.67 (Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Two or More Years Prior to Becoming an NBCT	0.0130 (0.0112)	0.0431** (0.0108)	0.0204* (0.0097)	0.0207* (0.0085)
Application Year for NBCT	0.0054 (0.0125)	0.0348** (0.0125)	0.0073 (0.0108)	0.0291** (0.0091)
Received Natl. Board Certification	0.0256** (0.0081)	0.0634** (0.0057)	0.0177** (0.0064)	0.0352** (0.0042)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	Yes	Yes	Yes	Yes
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	No	No	No	No
School Fixed Effects	No	No	No	No
R-squared	0.084	0.146	0.092	0.107
No. of Student-Year Obs.	454054	580934	477382	585432

Note: dependent variable is the normalized (by grade and year) achievement level on the test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina) minus 0.67 times the lagged test score. For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus in year 2000/01 lagged scores are available for grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus lagged scores are available in year 2000/01 in elementary school for North Carolina in grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 9
Estimates of the Effects of NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina -- Baseline Model Less Student FE, Plus Student Demographics, Plus School FE with Variable Persistence in Prior Schooling Inputs (Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Two or More Years Prior to Becoming an NBCT	-0.0017 (0.0107)	0.0346** (0.0103)	0.0162+ (0.0095)	0.0128 (0.0080)
Application Year for NBCT	-0.0110 (0.0123)	0.0329** (0.0116)	0.0066 (0.0104)	0.0215* (0.0086)
Received Natl. Board Certification	0.0179* (0.0080)	0.0529** (0.0055)	0.0185** (0.0065)	0.0269** (0.0042)
Lagged Test Score	0.6670** (0.0015)	0.7122** (0.0011)	0.6618** (0.0015)	0.6756** (0.0012)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	Yes	Yes	Yes	Yes
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	No	No	No	No
School Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.658	0.732	0.641	0.672
No. of Student-Year Obs.	454054	580934	477382	585432

Note: dependent variable is the normalized (by grade and year) achievement level on the test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina). The lagged test score is included as a regressor. For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus in year 2000/01 lagged scores are available for grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus lagged scores are available in year 2000/01 in elementary school for North Carolina in grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 10
Estimates of the Effects of NBPTS-Certified Teachers On Student Achievement in Florida and North Carolina -- Baseline Model Less Student FE, Plus Student Demographics, Plus School FE with Persistence in Prior Schooling Inputs Fixed at 0.67 (Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	FL	NC	FL	NC
Two or More Years Prior to Becoming an NBCT	-0.0017 (0.0108)	0.0336** (0.0103)	0.0164+ (0.0095)	0.0131+ (0.0079)
Application Year for NBCT	-0.0111 (0.0123)	0.0335** (0.0116)	0.0066 (0.0104)	0.0214* (0.0087)
Received Natl. Board Certification	0.0178* (0.0080)	0.0532** (0.0055)	0.0184** (0.0065)	0.0270** (0.0042)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	Yes	Yes	Yes	Yes
Classroom Characteristics	Yes	Yes	Yes	Yes
Student Fixed Effects	No	No	No	No
School Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.114	0.186	0.107	0.124
No. of Student-Year Obs.	454054	580934	477382	585432

Note: dependent variable is the normalized (by grade and year) achievement level on the test used for accountability in each state (the "Sunshine State Standards" test in Florida and the "End-of-Grade" test in North Carolina) minus 0.67 times the lagged test score. For Florida, in year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus in year 2000/01 lagged scores are available for grades 5 and 9 in reading and grades 6 and 9 in math. All grades 3-8 were tested in year 1999/00 in North Carolina, thus lagged scores are available in year 2000/01 in elementary school for North Carolina in grades 4 and 5 in both math and reading. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. For North Carolina the reported number of observations includes all students with relevant data. For Florida the reported number of observations includes only students with at least three valid test scores.

Table 11
"Baseline" Model
Estimates of the Effects of NBPTS-Certified
Teachers on Student Achievement in Florida by Grade Level
(Grades 4-10, 2000/01-2003/04)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High (Grades 9-10)
Two or More Years Prior to Becoming an NBCT	-0.0069 (0.0147)	0.0497** (0.0140)	0.0051 (0.0106)	0.0462** (0.0143)	0.0629** (0.0116)	-0.0119 (0.0124)
Application Year for NBCT	-0.0053 (0.0169)	-0.0253 (0.0155)	-0.0007 (0.0114)	-0.0095 (0.0156)	-0.0008 (0.0112)	-0.0588** (0.0128)
Achieved NBCT Status	0.0107 (0.0106)	-0.0521** (0.0139)	0.0266** (0.0086)	0.0091 (0.0093)	0.0227** (0.0072)	0.0048 (0.0098)
Student Time-Varying Char.	Yes	Yes	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Time-Varying Char.	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No	No	No
R-squared	0.611	0.618	0.610	0.668	0.649	0.585
No. of Student-Year Obs.	478599	436938	391117	620886	587307	449287

Note: dependent variable is the gain in the normalized (by grade and year) "Sunshine State Standards" achievement test score. In year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus year 2000/01 gains include grades 5 and 9 in reading and grades 6 and 9 in math. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. Samples include only those students with at least three test scores.

Table 12
"Baseline" Model Plus School Fixed Effects
Estimates of the Effects of NBPTS-Certified
Teachers on Student Achievement in Florida by Grade Level
(Grades 4-10, 2000/01-2003/04)

	Math			Reading		
	Elementary (Grades 4-5)	Middle (Grades 6-8)	High (Grades 9-10)	Elementary (Grades 4-5)	Middle (Grades 6-8)	High (Grades 9-10)
Two or More Years Prior to Becoming an NBCT	-0.0099 (0.0153)	0.0565** (0.0146)	-0.0040 (0.0108)	0.0434** (0.0151)	0.0690** (0.0124)	-0.0191 (0.0128)
Application Year for NBCT	0.0051 (0.0177)	-0.0291+ (0.0163)	-0.0003 (0.0116)	-0.0086 (0.0165)	-0.0035 (0.0116)	-0.0697** (0.0131)
Achieved NBCT Status	0.0217+ (0.0112)	-0.0617** (0.0147)	0.0224* (0.0092)	0.0089 (0.0100)	0.0244** (0.0075)	0.0051 (0.0102)
Student Time-Varying Char.	Yes	Yes	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Time-Varying Char.	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.670	0.659	0.639	0.718	0.684	0.615
No. of Student-Year Obs.	478599	436938	391117	620886	587307	449287

Note: dependent variable is the gain in the normalized (by grade and year) "Sunshine State Standards" achievement test score. In year 1999/00 reading was tested in grades 4, 8 and 10 and math was tested in grades 5, 8 and 10. Thus year 2000/01 gains include grades 5 and 9 in reading and grades 6 and 9 in math. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. Samples include only those students with at least three test scores.

Table 13
Estimates of the Effects of NBPTS-Certified Teachers on Student Achievement in Florida by Exam Type (Grades 4-5, 2000/01-2003/04)

	Math		Reading	
	SSS Test	Stanford-9	SSS Test	Stanford-9
Two or More Years Prior to Becoming an NBCT	-0.0069	0.0391**	0.0462**	0.0222*
	(0.0147)	(0.0103)	(0.0143)	(0.0094)
Application Year for NBCT	-0.0053	-0.0239+	-0.0095	-0.0035
	(0.0169)	(0.0138)	(0.0156)	(0.0130)
Achieved NBCT Status	0.0107	0.0055	0.0091	-0.0115
	(0.0106)	(0.0091)	(0.0093)	(0.0078)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes
Teacher Time-Varying Char.	Yes	Yes	Yes	Yes
Teacher Demographics	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No
R-squared	0.611	0.537	0.668	0.532
No. of Student-Year Obs.	478599	727630	620886	738943

Note: dependent variable is the gain in the normalized (by grade and year) achievement test score. In year 1999/00 the Sunshine State Standards exam was administered only in grades 4, 8 and 10 for reading and grades 5, 8 and 10 for math. Thus year 2000/01 gains include grades 5 and 9 in reading and grades 6 and 9 in math. The Stanford-9 exam was administered in all grades 3-10 in school year 1999/2000. Thus year 2000/01 gains for the Stanford-9 include all grades 4-10 for both reading and math. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. Samples include only those students with at least three test scores.

Table 14
Estimates of the Effects of NBPTS-Certified Teachers on Student Achievement in Florida by Exam Type (Grades 6-8, 2000/01-2003/04)

	Math		Reading	
	SSS Test	Stanford-9	SSS Test	Stanford-9
Two or More Years Prior to Becoming an NBCT	0.0497**	0.0548**	0.0629**	0.0171*
	(0.0140)	(0.0125)	(0.0116)	(0.0083)
Application Year for NBCT	-0.0253	-0.0190	-0.0008	-0.0019
	(0.0155)	(0.0162)	(0.0112)	(0.0103)
Achieved NBCT Status	-0.0521**	-0.0398**	0.0227**	0.0123+
	(0.0139)	(0.0131)	(0.0072)	(0.0072)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes
Teacher Time-Varying Char.	Yes	Yes	Yes	Yes
Teacher Demographics	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No
R-squared	0.618	0.623	0.649	0.555
No. of Student-Year Obs.	436938	522570	587307	783022

Note: dependent variable is the gain in the normalized (by grade and year) achievement test score. In year 1999/00 the Sunshine State Standards exam was administered only in grades 4, 8 and 10 for reading and grades 5, 8 and 10 for math. Thus year 2000/01 gains include grades 5 and 9 in reading and grades 6 and 9 in math. The Stanford-9 exam was administered in all grades 3-10 in school year 1999/2000. Thus year 2000/01 gains for the Stanford-9 include all grades 4-10 for both reading and math. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. Samples include only those students with at least three test scores.

Table 15
Estimates of the Effects of NBPTS-Certified Teachers on Student Achievement in Florida by Exam Type (Grades 9-10, 2000/01-2003/04)

	Math		Reading	
	SSS Test	Stanford-9	SSS Test	Stanford-9
Two or More Years Prior to Becoming an NBCT	0.0051 (0.0106)	0.0029 (0.0126)	-0.0119 (0.0124)	0.0198+ (0.0116)
Application Year for NBCT	-0.0007 (0.0114)	0.0055 (0.0161)	-0.0588** (0.0128)	-0.0265+ (0.0138)
Achieved NBCT Status	0.0266** (0.0086)	0.0243* (0.0108)	0.0048 (0.0098)	-0.0035 (0.0102)
Student Time-Varying Char.	Yes	Yes	Yes	Yes
Student Demographics	No	No	No	No
Classroom Characteristics	Yes	Yes	Yes	Yes
Teacher Time-Varying Char.	Yes	Yes	Yes	Yes
Teacher Demographics	Yes	Yes	Yes	Yes
Student Fixed Effects	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No
R-squared	0.610	0.597	0.585	0.590
No. of Student-Year Obs.	391117	382859	449287	437190

Note: dependent variable is the gain in the normalized (by grade and year) achievement test score. In year 1999/00 the Sunshine State Standards exam was administered only in grades 4, 8 and 10 for reading and grades 5, 8 and 10 for math. Thus year 2000/01 gains include grades 5 and 9 in reading and grades 6 and 9 in math. The Stanford-9 exam was administered in all grades 3-10 in school year 1999/2000. Thus year 2000/01 gains for the Stanford-9 include all grades 4-10 for both reading and math. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test. Included time-varying student characteristics are indicators for repeating a grade, "structural" move from another school and "non-structural" move from another school. Student demographics include race, ethnicity, gender, limited-English proficiency, "gifted" status, disability status, and free/reduced-price lunch receipt. Elementary regressions also include the student's age the first time in grade 3. Classroom characteristics include class size and the proportion of classmates who are non-white. Teacher time-varying characteristics include six experience categories (1-2, 3-5, 6-12, 13-20, 21-27, 28+ years), possession of an advanced degree, and possession of a standard teaching certificate. Teacher demographics include race, ethnicity, gender and indicators for matches between the race/ethnicity and gender of the teacher and the student. Samples include only those students with at least three test scores.

Table 16
Estimates of the Effects of NBPTS-Certified Teachers on Student Math Achievement in North Carolina (Grades 4-5)

	99/00 – 03/04		96/97 – 03/04	
	Basic Model	with Added Teacher Var.	Basic Model	with Added Teacher Var.
Two or More Years Prior to Becoming an NBCT	0.0544** (0.0119)	0.0522** (0.0119)	0.0410** (0.0051)	0.0389** (0.0051)
Application Year for NBCT	0.0268* (0.0125)	0.0245* (0.0125)	0.0470** (0.0084)	0.0447** (0.0084)
Achieved NBCT Status	0.0588** (0.0060)	0.0566** (0.0061)	0.0439** (0.0050)	0.0412** (0.0050)
Teacher Average Exam Score		0.0090** (0.0021)		0.0124** (0.0013)
Undergrad. Inst. – Very Competitive		0.0184 (0.0187)		-0.0032 (0.0151)
Undergrad Inst. - Competitive		0.0222 (0.0185)		0.0004 (0.0150)
Undergrad Inst. – Less Competitive		-0.0004 (0.0186)		-0.0126 (0.0151)
Undergrad. Inst. – Not Rated		0.0134 (0.0197)		-0.0047 (0.0158)
R-squared	0.557	0.557	0.431	0.431
No. of Student-Year Obs.	757111	757111	1501705	1501705

Note: dependent variable is the gain in the normalized (by grade and year) achievement test score. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test.

Table 17
Estimates of the Effects of NBPTS-Certified Teachers on Student Reading Achievement in North Carolina (Grades 4-5)

	99/00 – 03/04		96/97 – 03/04	
	Basic Model	with Added Teacher Var.	Basic Model	with Added Teacher Var.
Two or More Years Prior to Becoming an NBCT	0.0344** (0.0100)	0.0339** (0.0100)	0.0046 (0.0039)	0.0039 (0.0039)
Application Year for NBCT	0.0197+ (0.0103)	0.0192+ (0.0103)	0.0263** (0.0065)	0.0257** (0.0065)
Achieved NBCT Status	0.0276** (0.0051)	0.0274** (0.0051)	0.0191** (0.0040)	0.0185** (0.0040)
Teacher Average Exam Score		0.0018 (0.0018)		0.0044** (0.0011)
Undergrad. Inst. – Very Competitive		-0.0475** (0.0161)		-0.0331* (0.0131)
Undergrad Inst. - Competitive		-0.0392* (0.0159)		-0.0268* (0.0130)
Undergrad Inst. – Less Competitive		-0.0496** (0.0161)		-0.0323* (0.0131)
Undergrad. Inst. – Not Rated		-0.0495** (0.0170)		-0.0371** (0.0137)
R-squared	0.545	0.545	0.419	0.419
No. of Student-Year Obs.	769103	769103	1536297	1536297

Note: dependent variable is the gain in the normalized (by grade and year) achievement test score. Absolute values of robust standard errors (clustered at the classroom level) appear in parentheses. A + indicates statistical significance at the .10 level, * indicates significance at the .05 level and ** indicates significance at the .01 level in a two-tailed test.