

A Degree Above? The Value-Added Estimates and Evaluation Ratings of Teachers with a Graduate Degree

Kevin C. Bastian
Senior Research Associate
Associate Director
Education Policy Initiative at Carolina
University of North Carolina at Chapel Hill
314 Cloister Court
Chapel Hill, NC 27514
kbastian@email.unc.edu

Running Head: Graduate degrees and teacher performance

Manuscript submitted to *Education Finance and Policy* on January 30, 2017; resubmitted on August 3, 2017, December 14, 2017, and January 30, 2018

Acknowledgements

I wish to thank the University of North Carolina General Administration for its continued promotion and support of the Teacher Quality Research Initiative and the deans of the Schools and Colleges of Education at the 15 UNC system institutions engaged in teacher education for their comments and feedback throughout the research process. I am also very appreciative of two anonymous reviewers and the editorial team at *Education Finance and Policy* for their beneficial feedback. I acknowledge funding for this project from the Teacher Quality Research Initiative.

Abstract

In the present study I use teacher value-added and evaluation rating data from North Carolina public schools to estimate the signaling and human capital effects of graduate degrees. These analyses consider the effects of graduate degrees, overall, and the effects of graduate degrees inside and outside teachers' area(s) of teaching. Signaling analyses show that those with a graduate in their area of teaching have comparable value-added estimates and receive higher evaluation ratings than teachers with undergraduate degrees only. Human capital analyses indicate that in-area graduate degrees benefit teacher value-added in several comparisons and predict higher evaluation ratings on the Leadership standard. Signaling and human capital effects for out-of-area graduate degrees are generally negative or insignificant. Taken together, these results present a more comprehensive and nuanced view of the effectiveness of teachers with graduate degrees. Future analyses should assess additional outcome measures and continue focusing on the alignment between the graduate degree content and the teaching assignment.

<A> 1. Introduction

Given the relationships between teachers and short and long-term student outcomes (Chetty, Friedman, and Rockoff, 2014a, 2014b), states and school districts frequently seek approaches to reward effective teachers. While education officials and practitioners have instituted many policies and reforms towards this end—e.g. career ladders, pay for performance—one long-standing approach has been to incentivize teachers to acquire additional credentials. Most notably, many states and school districts provide permanent salary increases for teachers that earn a graduate degree (Roza and Miller, 2009). For example, Roza and Miller (2009) report that prior to the Great Recession policymakers in every state provided financial incentives to teachers who obtained graduate degrees. As of 2007, these graduate degree salary increases were estimated to be 2.1 percent of all K-12 education expenditures, with the average salary increase for graduate degrees costing \$174 per pupil.

In the educator labor market these graduate degrees can serve two main purposes: (1) as a *signal* of quality—that teachers with a graduate degree are more effective than those without the credential—and (2) as an indicator of *human capital* development—that the process of earning a graduate degree enhanced teachers' knowledge and skills. Regarding the signaling effect of graduate degrees, research generally shows that graduate degree holders are no more effective than teachers with undergraduate degrees only (Clotfelter, Ladd, and Vigdor, 2006, 2007, 2010; Buddin and Zamarro, 2009; Chingos and Peterson, 2011). Likewise, assessments of human capital development show that the process of earning a graduate degree does not consistently boost teacher effectiveness (Chingos and Peterson, 2011; Harris and Sass, 2011; Ladd and Sorensen, 2015). Given these research findings, pay increases for graduate degree

holders have faced critical attention in recent years, with a small number of districts and states limiting or ending graduate degree salary supplements.

While these decisions to curtail graduate degree pay may be viewed as evidence-based policymaking, there are several concerns with the extant research on graduate degree impacts. First, most prior studies have only considered the overall effect of graduate degrees and have not assessed whether the content of the graduate degree is associated with teacher effectiveness. Second, previous analyses may not have correctly modeled the human capital effect since potential gains in teacher knowledge and skills can accrue while a teacher is enrolled in graduate degree coursework. Finally, most prior studies have not considered whether graduate degrees have a signaling or human capital effect on outcomes beyond value-added. In the present study I address these concerns by coding the content of graduate degrees, estimating models that better assess human capital effects, and considering multiple measures of teacher effectiveness.

Specifically, I answer two policy relevant questions:

- (1) Are teachers with a graduate degree more effective than their peers with undergraduate degrees only?
- (2) Are teachers more effective after earning a graduate degree?

To answer these questions, I leverage administrative data from North Carolina public schools (NCPS) that allow me to code the content and timing of graduate degrees and to assess teacher value-added and teacher evaluation ratings. I also report value-added estimates and evaluation ratings for Nationally Board Certified (NBC) teachers. This lets me compare the signaling and human capital effects of graduate degrees with another teacher credential typically associated with teacher salary increases. Overall, this work contributes to policy discussions on

teacher salary supplements and fits with current research on the multi-dimensionality of teacher effectiveness and the benefits of job-specific human capital.

In the remaining sections I provide background on teacher pay in North Carolina and review the literature on graduate degree and NBC outcomes. Next, I describe the research sample, how I code graduate degrees, the measures used in analyses, and my methodological approaches. Finally, I present the results and close with a discussion of policy implications.

<A> 2. Background

 Teacher Credentials and Pay Schedules in North Carolina

The 115 public school districts in North Carolina compensate teachers according to a single, statewide salary structure, with districts able to differentiate teacher pay with local salary supplements.¹ Like many states and school districts, North Carolina's salary structure provides additional pay to more experienced teachers, teachers with graduate degrees, and teachers with NBC. Specifically, beginning in the 1997-98 school year, North Carolina instituted 12 percent pay increases for NBC teachers; shortly thereafter, in the 2000-01 school year, the state also instituted 10, 15, and 20 percent pay increases for teachers with masters, specialist, or doctoral degrees.² By 2013-14, approximately 12 and 36 percent of the state's 97,000 teachers were NBC or held a graduate degree. North Carolina leads the nation in its percentage of teachers with NBC but employs a smaller percentage of teachers with graduate degrees than many other states. As such, North Carolina apportions a smaller percentage of its total education budget to graduate degree salary supplements than the national average (Roza and Miller, 2009).

¹ In 2013-14 the average teacher salary supplement in North Carolina was \$3,553. Average salary supplements in districts ranged from \$0 (8 school districts did not offer a salary supplement) to over \$6,000 in Charlotte-Mecklenburg, Chapel Hill-Carrboro, and Wake County. Additionally, a small number of school districts are now experimenting with differentiated teacher pay based on leadership roles and performance.

² Prior to 2000-01, the pay increases for masters, specialist, and doctoral degrees in North Carolina were approximately 6.20, 11.50, and 16.70 percent, respectively.

In response to the budgetary pressures of the Great Recession and the research evidence on graduate degree impacts, the North Carolina General Assembly eliminated pay increases for graduate degree attainment in 2013 (Clotfelter, Ladd, and Vigdor, 2007, 2010; Buddin and Zamarro, 2009; Chingos and Peterson, 2011). Teachers who had already earned a graduate degree or those who had completed at least one graduate-level course by August 2013 would continue to receive or be eligible to receive salary increases. All other graduate degrees would not result in salary increases. These reforms to graduate degree pay were part of a larger trend in teacher compensation policy in North Carolina. Prior to the Great Recession, average teacher salaries in the state ranked 25th nationally; as of 2014-15, the state's average teacher salary ranked 42nd (National Education Association, 2010, 2014).

 Prior Literature

With the rise of test-based accountability systems and salary structures linked to teacher credentials, research efforts over the last two decades have examined the associations between teacher value-added and teachers' graduate degree and NBC status. Regarding graduate degrees, study findings generally show that the value-added of graduate degree holders is similar to that of teachers with undergraduate degrees only (Rockoff, 2004; Jepsen, 2005; Hanushek, Kain, O'Brien, and Rivkin, 2005; Clotfelter, Ladd, and Vigdor, 2006, 2007, 2010; Buddin and Zamarro, 2009; Chingos and Peterson, 2011). There are exceptions to these results, as Betts and colleagues (2003), Dee (2004), and Nye and colleagues (2004) return positive signaling effects in elementary grades mathematics. Conversely, Rockoff (2004) and Clotfelter and colleagues (2006, 2007) find negative signaling results in elementary grades reading. Likewise, research to assess whether graduate degrees develop the human capital of teachers returns few statistically significant value-added results. Using data from Florida, Harris and Sass (2011) find positive

human capital effects in middle grades mathematics but negative effects in middle grades reading. With data from North Carolina's high school End-of-Course (EOC) exams, Ladd and Sorensen (2015) return negative human capital effects in algebra 2 and biology and a positive effect in physical science. Coupled with meta-analyses indicating that teachers with graduate degrees are less likely to remain in teaching (Borman and Dowling, 2008), this body of work challenges the efficacy of salary supplements for graduate degree holders.

Contrary to the tenor of these graduate degree results, prior research returns positive findings for NBC teachers. Specifically, signaling analyses generally show that NBC teachers are more effective than their non-NBC peers across elementary and high school grades (Clotfelter, Ladd, and Vigdor, 2007, 2010; Goldhaber and Anthony, 2007; Cowan and Goldhaber, 2016). One exception to this is Harris and Sass (2009), who find only isolated signaling results with student achievement data from Florida. Regarding human capital development, there is little evidence that teachers are more effective after completing the NBC process than they were prior to certification (Goldhaber and Anthony, 2007; Harris and Sass, 2009). This suggests that teachers who pursue NBC are already more effective than their non-NBC peers.

Building from this previous work, the present study makes several contributions to the research base on graduate degrees and NBC. First, I analyze two teacher performance outcomes: value-added estimates and evaluation ratings. This broader focus recognizes that teacher effectiveness is a multi-dimensional construct (Jennings and DiPrete, 2010; Blazar and Kraft, 2017) and that teachers with graduate degrees may impact other desired outcomes. For example, Ladd and Sorensen (2015) find that middle school teachers with a graduate degree are associated with reduced student absentee rates. For the present study, it is possible that graduate degree holders have higher quality instructional practices or make contributions to their schools that do

not translate into higher value-added but do show up in their evaluation ratings. Conversely, current research suggests that principals give higher evaluation ratings to teachers who they perceive as harder working and actively engaged in professional development (Harris, Ingle, and Rutledge, 2014). Teachers pursuing a graduate degree may fall into this category and as such, higher evaluation ratings for graduate degree holders may be attributable to their effort rather than their teaching skill or unique contributions to the school.

Second, I code the content of teachers' graduate degrees and am able to assess the signaling and human capital effects of graduate degrees that are inside and outside a teacher's area of teaching. These analyses may particularly benefit the policy discussions of states and school districts that are considering providing or re-instating graduate degree pay for in-area graduate degrees. Furthermore, these analyses fit with current work showing the benefits of job-specific human capital. Specifically, studies by Ost (2014), Blazar (2015), and Atteberry and colleagues (2017) show that teachers consistently assigned to the same grade-level or subject-area are more effective and develop more rapidly on-the-job. Teachers earning a graduate degree in their area of teaching may also acquire job-specific human capital that benefits their performance. Early analyses suggested that teachers with a master's degree in mathematics were more effective in that subject-area than peers without a graduate degree (Goldhaber and Brewer, 1996, 2000); recent analyses by Ladd and Sorensen (2015) reveal few significant effects between the content of the graduate degree and improvements in teacher value-added.

Third, I estimate models to more accurately identify any human capital effects of a graduate degree. Previous work has used a teacher fixed effect to compare changes in teacher effectiveness before and after graduate degree conferral (Chingos and Peterson, 2011; Harris and Sass, 2011; Ladd and Sorensen, 2015). However, because teachers typically take graduate degree

courses over multiple years, any benefits to obtaining a graduate degree may already be actualized by the time their course of study is complete. Any positive human capital effects may also be attributable to changes in teachers' obligations—after completing their graduate degree, teachers should have more time to devote to teaching responsibilities.

Finally, I clearly describe and estimate the signaling and human capital effects of graduate degrees. Prior studies have not fully detailed the rationale of each model and what their coefficient estimates represent. For example, researchers often present teacher fixed effect models as a way to adjust for teacher selection but do not make clear that these results identify human capital rather than signaling effects (Ladd and Sorensen, 2015). Likewise, previous analyses have generally only estimated one of these two types of models—e.g. using a student (Clotfelter, Ladd, and Vigdor, 2007, 2010) or school fixed effect (Buddin and Zamarro, 2009; Chingos and Peterson, 2011) for signaling results or a teacher fixed effect to assess human capital development (Harris and Sass, 2011; Ladd and Sorensen, 2015).

<A> 3. Data

 Research Sample

In the present study I focus on teachers in NCPS during the 2005-06 through 2013-14 academic years. For teacher value-added, the analysis sample includes all those teaching a tested grade or subject-area in the 2005-06 through 2013-14 school years. In analyses, I separate the value-added sample into 10 grade-level/subject-area combinations: three models for End-of-Grade (EOG) exams in elementary grades (mathematics, reading, and 5th grade science); three models for EOG exams in middle grades (mathematics, reading, and 8th grade science); and four models for EOC exams in high school grades (mathematics, English, science, and social studies). The specific subject-areas for these EOC exams are as follows: mathematics (algebra 1,

geometry, and algebra 2), English (English 1 and English 2), science (biology, chemistry, physical science, and physics), and social studies (US history and civics). For EOG exams, student test scores are available for all analysis years in mathematics and reading; test scores for 5th and 8th grade science are available beginning in 2008-09 and 2007-08, respectively. For EOC exams, test scores from algebra 1 are available for all analysis years. For other EOC subject areas, test score availability differs since there are no test score data for high school science exams (biology, chemistry, physical science, and physics) in 2006-07 (due to test piloting) and because North Carolina has discontinued many EOC exams in recent years. Specifically, EOC test scores are available as follows: (1) biology for 2005-06 through 2013-14; (2) algebra 2, geometry, physical science, US history, and civics for 2005-06 through 2010-11; (3) physics and chemistry for 2005-06 through 2008-09; and (4) English 1 for 2005-06 through 2011-12 and English 2 for 2012-13 and 2013-14. In total, across all elementary, middle, and high school models, I analyzed 15,736,449 test scores and 157,094 unique teachers.

For teacher evaluation ratings, the analysis sample includes all NCPS teachers evaluated by their school principal in the 2011-12 through 2013-14 academic years. In comparison to value-added analyses, which are limited to tested-grade/subject-area teachers and only provide a single estimate of effectiveness, two major benefits of evaluation ratings are the expanded sample and information on specific aspects of teaching quality (Goldring et al., 2015). Evaluation ratings are available for more than 90 percent of the state's teachers; by comparison, student test score data are available for approximately 30 to 40 percent of the state's teachers. These evaluation ratings come from the North Carolina Educator Evaluation System (NCEES), an evaluation rubric in place across NCPS in which principals rate teachers on up to five professional teaching standards: Teachers Demonstrate Leadership (Standard 1), Teachers

Establish a Respectful Classroom Environment for a Diverse Group of Students (Standard 2), Teachers Know the Content They Teach (Standard 3), Teachers Facilitate Learning for Their Students (Standard 4), and Teachers Reflect on Their Practice (Standard 5).

Teachers with three or fewer years of consecutive employment are evaluated on all five professional teaching standards. The evaluation process for these teachers includes a pre-observation conference, three formal observations by the school principal (lasting at least 45 minutes each), one peer observation, post-observation conferences, the submission of teaching artifacts, and an end-of-year evaluation conference with the school principal. These novice teachers must earn ratings of proficient on all five standards to be eligible for a continuing license. Teachers with more than three years of consecutive employment may be evaluated, at the discretion of the school, on all five standards or on the Leadership and Facilitating Student Learning standards only. The evaluation process for more experienced teachers includes two or three observations by the school principal (lasting at least 20 minutes each), the submission of teaching artifacts, and an end-of-year evaluation conference. Principals do not generate evaluation ratings after each observation session; instead, principals synthesize the evidence generated by observations and artifacts and provide teachers with a set of summative evaluation ratings during their end-of-year conference. Principals rate teachers on each evaluation standard as either not demonstrated (level 1), developing (level 2), proficient (level 3), accomplished (level 4), or distinguished (level 5). While a large percentage teachers are rated at proficient or above, the distribution of evaluation ratings—further detailed in the outcome measures section—has sufficient variation for analyses. In total, I analyzed over 267,000 evaluation ratings for nearly 115,000 unique teachers in the 2011-12 through 2013-14 school years.

[Insert Table 1 about here]

Table 1 displays basic teacher and school characteristics for all teachers in the analysis sample and for those with and without a graduate degree. Overall, teachers average over 12 years of experience and approximately 32 and 11 percent of the teacher-year observations are for graduate degree holders and NBC teachers, respectively. Comparing teachers with a graduate degree to their peers without, Table 1 indicates that graduate degree holders are older, have more teaching experience, and are much more likely to be NBC. Regarding school characteristics, teachers work in schools where approximately 54 percent of students qualify for subsidized meals, 48 percent of students are a racial/ethnic minority, and 67 percent of the standardized assessments are passed (performance composite). School characteristics for graduate degree holders are similar to those for teachers without a graduate degree.

 Coding Graduate Degrees

For these analyses I code graduate degrees in four ways. First, using the personnel education file from the North Carolina Department of Public Instruction (NCDPI)—which includes fields for degree level and graduation date—I create a time-varying indicator identifying all teachers with a graduate degree in each of my analysis years (2005-06 through 2013-14).

Second, because simply comparing teacher effectiveness in the pre and post-graduate degree periods may not accurately capture human capital effects, I account for the timing of graduate degree conferral. Specifically, I create the following set of indicators: *two years before degree conferral*, *one year before degree conferral*, *year of degree conferral*, *one year after degree conferral*, and *two years after degree conferral*. This coding scheme can help me assess whether teachers' value-added increases or decreases in the years immediately preceding degree conferral, in the year of degree conferral, and in the years immediately after degree conferral.

Third, I classify graduate degrees according to their content. The NCDPI personnel education file does not capture the content area for teachers' degrees. However, NCDPI licensure files identify the licensure area(s) teachers hold (e.g. elementary education, middle grades science) and the level of degree (e.g. bachelors, masters, specialist, doctorate) associated with that licensure area. Using these two fields I classified teachers' graduate degrees into the following eight categories: elementary education, special education, reading and English language arts, mathematics, science, social studies, school administration, and other (e.g. counseling, social work, curriculum specialist, foreign language, arts, career-technical education, etc.). Approximately 14 percent of the teacher-year records that I identified as having a graduate degree did not have a licensure area associated with a graduate degree in the licensure data. I retained these individuals in analyses and created a dichotomous indicator for *unclassified* graduate degrees.

Finally, as an extension of this content area coding, I used the value-added sample to create indicators for whether a teacher's graduate degree is in her tested-grade/subject-area or outside her tested-grade/subject-area. The in-area classifications are as follows: (1) teachers in elementary mathematics, reading, and science are in-area with a graduate degree in elementary education or a graduate degree in mathematics, English/reading, or science, respectively; (2) teachers in middle grades mathematics, reading, and science are in-area with a graduate degree in mathematics, English/reading, or science, respectively; (3) 6th grade teachers in mathematics and reading are in-area with a graduate degree in elementary education;³ and (4) teachers in high school mathematics, English, science, and social studies are in-area with a graduate degree in mathematics, English/reading, science, or social studies, respectively.

³ In North Carolina, the elementary education license covers grades K-6. Therefore, I classify 6th grade teachers with an elementary education graduate degree as in-area.

[Insert Figure 1 about here]

For those with a graduate degree (32.25 percent of the teacher-year records), Figure 1 displays the percentage of teacher-year observations for those earning a graduate degree before and after entry into teaching and for the content area coding. Approximately two-thirds of these teacher-year records are graduate degrees post-entry into teaching; the remaining records are for graduate degrees earned prior to beginning teaching. Regarding the content coding, the top two categories are other and elementary education. Graduate degrees in school administration—resulting in a license to be a school principal—make up eight percent of these teacher-year records. Lastly, graduate degrees in STEM fields—mathematics and science—comprise the smallest percentages of these records.

** Outcome Measures**

The dependent variable for value-added analyses is students' test scores on the North Carolina EOG and EOC exams. I standardized all EOG exams within subject, grade, and year and all EOC exams within subject and year. For EOG exams in mathematics, reading, and science, I used standardized mathematics and reading scores from the previous year as the measure of prior student achievement. For all high school EOC exams, the measure of prior student achievement is standardized mathematics and reading scores from the eighth grade. With standardized student achievement as the outcome measure, results tables display coefficients for graduate degrees and NBC in student-level standard deviation units. To make these value-added results comparable to those for teacher evaluation ratings (discussed below), results tables also display the standard deviation of individual teacher value-added for each test score model (e.g.

elementary mathematics, middle grades reading).⁴ This allows readers to more readily compare the magnitude of test score and evaluation rating coefficients.

The dependent variable for the teacher evaluation analyses is ratings from each of the five teaching standards that make up the NCEES (Leadership, Classroom Environment, Content Knowledge, Facilitating Student Learning, and Reflecting on Practice). For each standard, principals rate teachers as either not demonstrated, developing, proficient, accomplished, or distinguished (1-5). I standardized all evaluation ratings within each teaching standard and year so that I could interpret coefficients for graduate degrees and NBC as a percent of a standard deviation in teacher evaluation ratings.⁵

While teacher evaluation systems offer several benefits in comparison to value-added estimates—larger analysis samples, opportunities to measure distinct teaching practices (Ronfeldt and Campbell, 2016; Bastian, Patterson, and Pan, 2017)—there are potential concerns regarding the lack of variation in ratings (Toch and Rothman, 2008; Weisberg et al., 2009; Kraft and Gilmour, 2017) and the strength of associations between value-added estimates and evaluation ratings (Jacob and Lefgren, 2008; Harris, Ingle, and Rutledge, 2014). In my sample, the average ratings on the five professional teaching standards range between 3.53 and 3.73 (between proficient and accomplished). Figure 2 displays the distribution of ratings for each of the five teaching standards. Approximately 98 percent of teachers are rated at proficient or above. For the Leadership, Classroom Environment, and Facilitating Student Learning standards the modal rating is accomplished; for Content Knowledge and Reflecting on Practice the modal

⁴ Using data from 2005-06 through 2013-14, I estimated individual teacher-year value-added estimates in elementary mathematics, reading, and science; middle grades mathematics, reading, and science; and high school mathematics, English, science, and social studies. With these estimates I report the standard deviation in teacher value-added.

⁵ Prior to this standardization, I excluded cases where a teacher was rated as not demonstrated due to ambiguity in the rating and the limited use of the rating. Not demonstrated can refer to situations in which the principal is unable to observe or rate the indicators for the standard and is only used for approximately 0.10 percent of the sample.

rating is proficient. Regarding associations with value-added, Henry and Guthrie (2015) find that concurrent correlations between each of the NCEES standards and individual teacher value-added estimates are approximately 0.20 and statistically significant. These correlations are similar to those from previous studies examining the relationships between teacher performance measures (Rockoff et al., 2012; Harris, Ingle, and Rutledge, 2014).

[Insert Figure 2 about here]

<A> 4. Analysis Plan

Signaling analyses indicate whether teachers with a graduate degree are more effective than those without the credential. In these analyses any effectiveness differences may be attributable to teacher selection into a graduate degree and/or the knowledge and skills gained during the graduate degree program. Human capital analyses adjust for these teacher selection concerns and reveal whether the process of earning a graduate degree enhances teachers' effectiveness. From a policy perspective, these analyses evaluate two distinct aspects of graduate degree salary supplements: do pay increases reward teachers who are more effective and/or teachers who improve their effectiveness?

For teacher value-added I estimate linear regression models controlling for a rich set of covariates. While NCEES ratings are ordinal in nature (developing, proficient, accomplished, and distinguished), I also estimate linear regression models for teachers' evaluation ratings (standardized for analyses). Coefficients from these analyses are easier to interpret than odds ratios from ordered logit models; results from the linear and ordered logit approaches are very similar. Since ordered logit analyses do not easily accommodate fixed effects (Riedl and Geishecker, 2014), these linear regression models also have the advantage of allowing school or teacher fixed effects for evaluation rating analyses that align with my signaling and human

capital modeling. For both teacher value-added and evaluation ratings analyses I cluster standard errors at the school-by-year level; alternate approaches clustering standard errors at the teacher level return similar standard error values.

To assess the signaling effect of graduate degrees, I estimate covariate adjustment regression models with a school fixed effect (signified by the δ_s in equations 1 and 2). I prefer this approach since it adjusts for the sorting of teachers into schools and the impact of unobserved, time-invariant school characteristics on graduate degree estimates (Lankford, Loeb, and Wyckoff, 2002; Kennedy 2010; Goldhaber, Lavery, and Theobald, 2015). These analyses compare the effectiveness of teachers with a graduate degree to their peers without a graduate degree working in the same schools. For teachers who remain in the same school before and after earning a graduate degree, these analyses also leverage variation in their graduate degree status over time. Appendix Table A.1 displays counts of the unique teachers contributing to estimates. The unique teacher counts are much higher for the signaling than human capital analyses. This indicates that most of the signaling effect is identified by variation in graduate degree status across teachers working in the same schools.

To assess the human capital effect of graduate degrees, I substitute a teacher fixed effect (signified by the δ_j in equations 1 and 2) for the school fixed effect. This specification accounts for the non-random selection of teachers into graduate degree status. Essentially, the teacher fixed effect relies on within-teacher changes in graduate degree status and allows me to estimate whether teachers' effectiveness is significantly different after earning a graduate degree than before earning a graduate degree. Since human capital analyses rely on variation in graduate degree status within teachers, some effects are identified on a relatively small sample (see Appendix Table A.1). Because the benefits of a graduate degree may accrue during, rather than

after, the coursework period, I also explore teacher fixed effect specifications that include a set of graduate degree timing indicators. These analyses indicate how teacher value-added changes in the school years immediately preceding degree conferral, in the year of degree conferral, and in the school years immediately after degree conferral.⁶ Overall, the equations for the value-added and evaluation rating models are as follows:

$$Y_{ijst} = \beta_0 Ach_{ijst-x} + \beta_1 GradDeg_{jt} + \beta_2 NBC_{jt} + \beta_3 X_{ijst} + \beta_4 Z_{jst} + \beta_5 W_{st} + \delta_s/\delta_j + \varepsilon_{ijst} \quad (1)$$

$$Rating_{jst} = \beta_1 GradDeg_{jt} + \beta_2 NBC_{jt} + \beta_3 Z_{jst} + \beta_4 W_{st} + \delta_s/\delta_j + \varepsilon_{jst} \quad (2)$$

In these models Y_{ijst} is the test score for student i with teacher j in school s at time t (equation one) and $Rating_{jst}$ is the evaluation rating (on a particular NCEES standard) for teacher j in school s at time t (equation two). $GradDeg_{jt}$ is a time-varying characteristic for teachers' graduate degree status. These focal indicators include having a graduate degree (overall), the content area of the graduate degree, and whether the graduate degree was in or outside the area of teaching. In results tables I focus on the overall and in/out-of-area analyses; I table results from the content area models in Appendix Tables A.2 and A.3. NBC_{jt} is a time-varying characteristic for teachers' NBC status. Controlling for this credential helps me to better isolate outcomes for graduate degree holders—since many teachers with graduate degrees also have NBC—and allows for a comparison of results for two credentials traditionally associated with teacher pay.

In addition to these focal teacher measures, my value-added analyses control for students' prior exam scores (Ach_{ijst-x} in equation one) and a set of individual student, classroom/teacher, and school covariates (X_{ijst} , Z_{jst} , and W_{st}). The student covariates adjust for time-invariant and

⁶ I do not estimate these timing analyses on teacher evaluation ratings since I have only three years of evaluation data.

varying characteristics and include the following: the average prior achievement scores of a student's peers, year fixed effects, and indicators for mobility, being underage or overage, giftedness, disability, economic-disadvantage, gender, race/ethnicity status, and limited English proficiency. The high school EOC value-added models also control for subject-area (e.g. algebra 2 and geometry in a mathematics model) and grade-level. These subject-area and grade-level fixed effects help adjust for the fact that not all students take each EOC exam and that students take EOC exams at different points in their high school tenure. The classroom/teacher characteristics account for items of classroom structure and teacher credentials that may influence teaching quality and include the following: class size, the heterogeneity of prior achievement in the class, curriculum level (advanced or remedial), teaching experience, and teaching out-of-field. Finally, the school covariates adjust for characteristics that may impact teachers' choices to work at a given school and their performance while teaching there. These school covariates include school size, total per-pupil expenditures, average teacher salary supplements, short-term suspension and violent acts rates, and the percentages of economically-disadvantaged and racial/ethnic minority students. In evaluation rating analyses I control for teacher experience, the same set of school characteristics as in my value-added models, and indicators for school level (in reference to elementary schools). These school characteristics are particularly important, since research shows that classroom and school context are significantly associated with evaluation ratings (Whitehurst, Chingos, and Lindquist, 2014; Steinberg and Garrett, 2016).

To test the robustness of the graduate degree results, I estimate several specification checks. For teacher evaluation ratings, Appendix Table A.4 displays results from signaling and human capital analyses limited to the sample of teachers with value-added estimates. These

analyses account for the changes in sample across teacher outcomes, as many more teachers have evaluation ratings than value-added estimates. Likewise, Appendix Table A.4 presents results from signaling analyses that control for individual teacher value-added estimates.⁷ These analyses may help adjust for potential biases in principals' ratings of teachers.

Finally, given the number of significance tests performed, I acknowledge the possibility of Type I errors. It is possible that a small number of significance tests will result in false positive results for graduate degrees. I address this concern by focusing on consistent patterns in the graduate degree estimates. That is, I give greater credence to a graduate degree result when it is aligned with theory (e.g. the benefits of job-specific human capital) or when it is part of a series of significant findings rather than an isolated estimate. In my preferred approach I do not make formal significance test corrections since these methods can severely reduce the likelihood of detecting an effect (Gelman, Hill, and Yajima, 2012). As a robustness check, I do apply false discovery rate (FDR) corrections to my signaling and human capital results.⁸ Connected to these concerns about Type I error and statistical significance is the idea of practical significance. Given the large sample in my teacher value-added and evaluation rating analyses, I acknowledge that I may detect statistically significant results that are small in magnitude. Therefore, in the

⁷ Using data from 2011-12 through 2013-14, I estimate individual teacher-year value-added estimates in elementary grades mathematics, reading, and science; middle grades mathematics, reading, and science; and high school algebra 1, biology, and English. I standardized these value-added estimates within subject-area (e.g. elementary mathematics) and year. When a teacher had value-added estimates in more than one subject-area, I averaged their standardized estimates to have a single value-added measure.

⁸ I adjusted for the possibility of Type I errors with a false discovery rate (FDR) approach. For the value-added results from signaling analyses (Table 2), 18 estimates remain statistically significant after the FDR correction while two estimates are no longer significant (in-area for elementary grades reading and out-of-area for high school social studies). For the evaluation rating results from signaling analyses (Table 3), all 12 estimates remain statistically significant after the FDR correction. For the value-added results from human capital analyses (Table 4), three estimates remain statistically significant after the FDR correction (in-area for middle grades math, in-area for high school science, and unclassified for 8th grade science) while four estimates are no longer statistically significant. Finally, for the evaluation rating results from human capital analyses (Table 6), neither of the estimates (for the Leadership evaluation standard) remain statistically significant after the FDR correction.

results and discussion sections, I comment on the size of findings and benchmark them against other teacher credential measures. This helps convey the meaningfulness of a graduate degree to teacher performance.

<A> 5. Results

 Graduate Degrees as a Signal of Teacher Effectiveness

Considering teacher value-added, the top panel of Table 2 indicates that graduate degrees do not signal more effective teachers. Instead, there are five comparisons in which graduate degree holders are significantly less effective than teachers with undergraduate degrees only: elementary grades reading, middle grades reading, high school mathematics, high school science, and high school social studies. These estimates are small in magnitude: ranging from 0.4 to 1.3 percent of a student-level standard deviation and from four to eight percent of a standard deviation in teacher-level value-added. For example, the coefficient of -0.004 in elementary grades reading represents four percent of the teacher-level standard deviation ($-0.004/0.103$). By comparison, Table 2 indicates that NBC is a strong and consistent signal of teacher effectiveness—NBC teachers are significantly more effective than their non-NBC peers across all 10 grade-level/subject-area comparisons. These NBC coefficients range from 0.7 to 5.9 percent of a student-level standard deviation and from seven to 38 percent of the standard deviation in teacher-level value-added.

[Insert Table 2 about here]

Delving into these value-added results reveals signaling differences between in-area and out-of-area graduate degrees. As shown in Table 2, middle and high school mathematics teachers with an in-area graduate degree are more effective than those with an undergraduate degree only. These in-area estimates of 0.012 and 0.013 equate to eight percent of a standard deviation in

teacher-level value-added and are consistent with previous analyses showing that teachers with graduate degrees in mathematics have higher value-added estimates in that subject-area (Goldhaber and Brewer, 2000). Conversely, those with an out-of-area graduate degree are significantly less effective than teachers with an undergraduate degree only in seven of ten comparisons. Content-area results in the top panel of Appendix Table A.2 indicate that these negative signaling results are concentrated in teachers with School Administration and Other graduate degrees. Lastly, there are five negative and significant value-added results for the teacher-year graduate degree records whose content I could not classify.

[Insert Table 3 about here]

Considering teacher evaluation ratings, the top panel of Table 3 indicates that graduate degrees are a positive signal of teacher performance. Across all five professional teaching standards, graduate degree holders earn significantly higher evaluation ratings than teachers with undergraduate degrees only. For example, teachers with graduate degrees earn evaluation ratings on the Facilitating Student Learning standard that are 5.5 percent of a standard deviation higher than peers with undergraduate degrees only. These evaluation rating estimates are modest—ranging from 4.7 to 7.1 percent of a standard deviation—and comparable in size to the positive value-added results for in-area graduate degrees in Table 2. Furthermore, these evaluation results show little differentiation in ratings across teaching standards. That is, relative to those with undergraduate degrees, graduate degree holders are not recognized as performing any teaching task particularly better than other tasks. The bottom panel of Table 3 indicates that NBC is a strong signal of evaluation performance, with NBC teachers earning ratings 30 to 40 percent of a standard deviation higher than their non-NBC peers. These NBC results are roughly six times larger in magnitude than the evaluation results for graduate degree holders.

For the sample of teachers teaching a tested grade/subject-area in 2011-12 through 2013-14, the middle panel of Table 3 presents evaluation rating results from in-area analyses. These results show that teachers with an in-area graduate degree earn significantly higher evaluation ratings than those with an undergraduate degree only across all five professional teaching standards. The magnitude of these results is largest for the Content Knowledge standard and smallest for the Classroom Environment standard. For instance, teachers with an in-area graduate degree earn ratings 10 percent of a standard deviation higher on the Content Knowledge standard but only 3.6 percent of a standard deviation higher on the Classroom Environment standard. This indicates that the benefits of an in-area degree may be larger for teaching practices more closely aligned to the content of the degree. In contrast, the signaling effect of out-of-area and unclassified graduate degrees on evaluation ratings is generally insignificant. Taken together, this suggests that the in-area evaluation rating results are largely responsible for the pattern of significant graduate degree findings in the overall analyses.

Comparing the value-added and evaluation rating results, there are consistencies and discrepancies in the signaling effect of graduate degrees. Relative to the out-of-area findings, the in-area results are more positive for both value-added and evaluation ratings. Conversely, findings from the overall analyses are quite different: negative results for value-added versus consistently positive results for evaluation ratings.⁹ These differences spotlight a potential concern with evaluation ratings—while it is possible that evaluation ratings capture distinct aspects of teaching effectiveness, these ratings may also be biased by principals' beliefs (Harris, Ingle, and Rutledge, 2014). As a check on this, I re-estimated the overall and in-area analyses

⁹ It is possible that discrepancies in the overall results across outcomes are due to differences in the analytic sample. However, the top panel of Appendix Table A.4 displays positive and significant signaling results for an evaluation rating sample limited to teachers with value-added data.

controlling for individual teacher-year value-added. As shown in the middle panel of Appendix Table A.4, controlling for a more objective measure of teacher effectiveness does not meaningfully change the positive evaluation results for graduate degree holders. For evaluation ratings a graduate degree, particularly an in-area degree, remains a modest signal of teacher effectiveness.

 Human Capital Effects of Graduate Degrees

While a key policy concern is whether graduate degree holders are more effective than teachers with an undergraduate degree only, an important, related question is whether the process of earning a graduate degree boosts teachers' effectiveness. Considering teacher value-added, the top panel of Table 4 reveals positive and significant human capital results in 5th grade science and in high school English. These estimates represent 15 and 24 percent of a standard deviation in teacher-level value-added, respectively. Human capital results for NBC are generally insignificant—except for a negative finding in high school science. Relative to the positive signaling results in Table 2, this suggests that those earning NBC were already more effective than their non-NBC peers prior to completing the certification process. This is consistent with previous findings on NBC effects (Goldhaber and Anthony, 2007; Harris and Sass, 2009). Regarding the in-area analyses, in four comparisons—5th grade science, middle grades mathematics, high school science, and high school social studies—teachers are more effective after earning an in-area graduate degree. These estimates range from 2.0 to 3.8 percent of a student-level standard deviation and from 14 to 25 percent of a standard deviation in teacher-level value-added. These findings suggest the potential of graduate degree courses as a way for teachers to acquire job-specific human capital, particularly in STEM fields. In no instances is an out-of-area graduate degree associated with increases in teacher effectiveness.

[Insert Table 4 about here]

A concern with these teacher fixed effects analyses is whether they accurately estimate human capital effects. Specifically, these analyses compare teacher effectiveness pre and post graduate degree conferral, however, any benefits to the graduate degree coursework may have already accrued by degree completion. Conversely, teacher effectiveness may suffer during the coursework period if it prevents teachers from allocating sufficient time and attention to their teaching responsibilities. Therefore, I insert a set of timing indicators into the teacher fixed effect analyses to assess whether teacher value-added changes in the years immediately preceding degree conferral, in the year of degree conferral, or in the years immediately after degree conferral. Results in Table 5 do not indicate any consistent patterns in the timing of human capital returns. There is little evidence that teachers are getting better during their coursework period; likewise, findings do not support the hypothesis that teachers are less effective during their coursework period. In middle and high school grades, some evidence suggests that potential human capital effects may take more than a year to develop.

[Insert Table 5 about here]

Table 6 addresses whether graduate degrees have a human capital effect on teacher evaluation ratings. These analyses are limited, as I only assess three years of evaluation data. A longer data panel, allowing more time for teachers to earn graduate degrees, may reveal different findings. These results suggest targeted human capital benefits to a graduate degree. Specifically, there are positive findings for the Leadership standard, both overall and for in-area degrees, and null results for the remaining teaching standards.¹⁰ These findings are robust to specifications

¹⁰ Appendix Table A.3 indicates that teachers with an Administration graduate degree earn higher evaluation ratings on the Leadership standard than their undergraduate degree only peers. Conversely, Appendix Table A.3 displays negative human capital effects for these teachers on the Leadership standard.

limited to teachers with value-added data (bottom of Appendix Table A.4). The Leadership standard is comprised of indicators for teachers leading in their classrooms, in the school, and in the teaching profession and in advocating for positive changes for schools and students. These positive results suggest that earning a graduate degree may enhance the leadership role of teachers in schools. Furthermore, these results suggest that any human capital effects of a graduate degree may extend beyond teaching practices that are directly associated with student achievement.

[Insert Table 6 about here]

<A> 6. Discussion

This study contributes to current knowledge on teacher effectiveness and compensation by assessing the content of graduate degrees and examining multiple measures of teacher performance. While salary supplements for all graduate degrees are not well-supported by extant research (including findings from this study), my analyses show that in-area graduate degrees are related to teacher effectiveness. These in-area findings are generally robust to FDR corrections (particularly for the signaling analyses), suggesting that Type I error does not seriously threaten this study's key implications. Teachers with in-area graduate degrees are more effective in middle and secondary grades mathematics and the process of earning an in-area degree boosts teacher value-added in multiple subject-areas. Furthermore, teachers with in-area graduate degrees earn higher evaluation ratings on all five professional teaching standards. These results fit with recent research showing the benefits of job-specific human capital and have several implications for policy and research (Ost, 2014; Blazar, 2015).

Most notably, these results indicate that salary increases for in-area graduate degrees may represent an evidence-based policy. States and districts that have eliminated graduate degree pay

may wish to re-instate it for teachers earning in-area degrees; those still incentivizing all graduate degrees may wish to limit their salary increases to in-area degrees. To more fully evaluate this policy implication, it is beneficial to consider the magnitude of my in-area estimates. For teacher value-added the positive results for in-area degrees are modest, typically less than three percent of a standard deviation in student achievement. To put these results into perspective, Ladd and Sorensen (2017) find that middle grades mathematics teachers gain seven percent of a standard deviation in student achievement between their first and second-year of teaching and 11 percent of a standard deviation by their fifth-year of teaching. In middle grades reading, these values are two percent and four percent, respectively. The evaluation rating results for in-area graduate degrees are also relatively small—approximately four to 10 percent of a standard deviation. By comparison, my own analyses show that second-year teachers earn evaluation ratings 25 percent of a standard deviation higher than first-year teachers and that fifth-year teachers earn ratings 80 percent of a standard deviation higher. The return on an in-area graduate degree is generally smaller than teachers' on-the-job productivity gains. Perhaps more relevantly, my signaling analyses show that the return on an in-area graduate degree is smaller than that for NBC, another credential associated with teacher salary supplements. Taken together, these points suggest that policymakers may need to consider statistical and practical significance when evaluating credential-based compensation policies.

Beyond the direct salary implications for in-area degrees, my analyses challenge policymakers to more clearly communicate the motivations for credential-based salary supplements. Compensation policies for teacher credentials can be a mechanism to reward more effective teachers (signaling) and/or encourage teachers to improve (human capital). Currently, the goals of these policies are not well-specified—it is unclear whether policymakers want to

reward more effective teachers and/or those who become more effective. As a result, it is challenging to convey the purpose of these compensation policies to teachers and to evaluate the success of these policies. If policymakers consider refinements to credential-based compensation, they should assess and clearly communicate the scope of funding—e.g. for what credentials, for what content areas—and the purpose of that funding.

Lastly, my analyses have implications for the direction of future research and its contributions to policy. With the proliferation of PK-12 administrative data, continued research needs to estimate the signaling and human capital effects of teacher credentials on a range of measures (e.g. rubric-based classroom observations, student surveys, student attendance, teacher leadership in the school). This will help educators and policymakers have a more complete understanding of the effects of graduate degrees and NBC. This is particularly important in a policy context that is beginning to expand its definition (beyond student achievement) of highly-effective teachers. In addition, future analyses should continue to focus on the alignment between the content of the teacher credential and the teaching assignment. Job-specific human capital makes up a meaningful portion of overall teacher effectiveness (Cook and Mansfield, 2016) and represents a way in which education officials and policymakers can enhance teaching quality. Researchers need to further understand what types of learning opportunities (e.g. graduate degrees, NBC) promote teacher development so that policy can promote and incentivize these activities.

<A> 7. References

- Atteberry, Allison, Susanna Loeb, and James Wyckoff. 2017. Teacher churning: Reassignment rates and implications for student achievement. *Educational Evaluation and Policy Analysis* 39(1): 3-30.
- Bastian, Kevin C., Kristina M. Patterson, and Yi Pan. 2017. Evaluating teacher preparation programs with teacher evaluation ratings: Implications for program accountability and improvement. *Journal of Teacher Education*. In press.
- Betts, Julian R., Andrew C. Zau, and Lorien A. Rice. 2003. Determinants of student achievement: New evidence from San Diego. Unpublished paper, Public Policy Institute of California.
- Blazar, David. 2015. Grade assignments and the teacher pipeline: A low-cost lever to improve student achievement. *Educational Researcher* 44(4): 213-227.
- Blazar, David and Matthew A. Kraft. 2017. Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis* 39(1): 146-170.
- Borman, Geoffrey D. and N. Maritza Dowling. 2008. Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research* 78(3): 367-409.
- Buddin, Richard and Gema Zamarro. 2009. Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics* 66(2): 103-115.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *The American Economic Review* 104(9): 2593-2632.

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

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

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. Teacher student matching and the assessment of teacher effectiveness. *Journal of Human Resources* 41(4): 778-820.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2007. Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review* 26(6): 673-682.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources* 45(3): 655-681.

Dee, Thomas S. 2004. Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics* 86(1): 195-210.

Cook, Jason B. and Richard K. Mansfield. 2016. Task-specific experience and task-specific talent: Decomposing the productivity of high school teachers. *Journal of Public Economics* 140(8): 51-72.

Gelman, Andrew, Jennifer Hill, and Masanao Yajima. 2012. Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness* 5(2): 189-211.

Goldhaber, Dan D. and Dominic J. Brewer. 1996. Evaluating the effect of teacher degree level on educational performance. Available <http://files.eric.ed.gov/fulltext/ED406400.pdf>. Accessed July 13, 2016.

Goldhaber, Dan D. and Dominic J. Brewer. 2000. Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis* 22(2): 129-145.

Goldhaber, Dan, Lesley Lavery, and Roddy Theobald. 2015. Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher* 44(5): 293-307.

Goldring, Ellen, Jason A. Grissom, Mollie Rubin, Christine M. Neumerski, Marisa Cannata, Timothy Drake, and Patrick Schuermann. 2015. Make room value added: Principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher* 44(2): 96-104.

Hanushek, Eric A., John F. Kain, Daniel M. O'Brien, and Steven G. Rivkin. 2005. The market for teacher quality. NBER Working Paper No. 11154.

Harris, Douglas N. and Tim R. Sass. 2011. Teacher training, teacher quality, and student achievement. *Journal of Public Economics* 95(7): 798-812.

Harris, Douglas N., William K. Ingle, and Stacey A. Rutledge. 2014. How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal* 51(1): 73-112.

- Henry, Gary T. and James E. Guthrie. 2015. Using multiple measures of developmental teacher evaluation. In *Improving Teacher Evaluation Systems: Making the Most of Multiple Measures*, edited by Jason Grissom and Peter Youngs, pp. 143-155. New York: Teachers College Press.
- Jacob, Brian A. and Lars Lefgren. 2008. Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics* 26(1): 101-136.
- Jennings, Jennifer L. and Thomas A. DiPrete. 2010. Teacher effects on social and behavioral skills in early elementary school. *Sociology of Education* 83(2): 135-159.
- Jepsen, Christopher. 2005. Teacher characteristics and student achievement: Evidence from teacher surveys. *Journal of Urban Economics* 57(2): 302-319.
- Kennedy, Mary M. 2010. Attribution error and the quest for teacher quality. *Educational Researcher* 39(8): 591-598.
- Kraft, Matthew A. and Allison F. Gilmour. 2017. Revisiting the Widget Effect: Teacher evaluation reforms and the distribution of teacher effectiveness. *Educational Researcher* 46(5): 234-249.
- Ladd, Helen F. and Lucy C. Sorensen. 2015. Do master's degrees matter? Advanced degrees, career paths, and the effectiveness of teachers. Calder Working Paper No. 136.
- Ladd, Helen F. and Lucy C. Sorensen. 2017. Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy* 12(2): 241-279.
- Lankford, Hamilton, Susanna Loeb, and James Wyckoff. 2002. Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis* 24(1): 37-62.
- National Education Association. *Rankings and estimates: Rankings of the states 2009 and estimates of school statistics 2010*. Available <http://www.nea.org/assets/docs/010rankings.pdf>. Accessed July 10, 2016.

National Education Association. *Rankings and estimates: Rankings of the states 2013 and estimates of school statistics 2014*. Available <http://www.nea.org/assets/docs/NEA-Rankings-and-Estimates-2013-2014.pdf>. Accessed July 10, 2016.

Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges. 2004. How large are teacher effects? *Educational Evaluation and Policy Analysis* 26(3): 237-257.

Ost, Ben. 2014. How do teachers improve? The relative importance of specific and general human capital. *American Economic Journal: Applied Economics* 6(2): 127–51.

Riedl, Maximilian and Ingo Geishecker. 2014. Keep it simple: Estimation strategies for ordered response models with fixed effects. *Journal of Applied Statistics* 41(11): 2358-2374.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. Teachers, schools, and achievement. *Econometrica* 73(2): 417-458.

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

Rockoff, Jonah E., Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. 2012. Information and employee evaluation: Evidence from a randomized intervention in public schools. *The American Economic Review* 102(7): 3184-3213.

Ronfeldt, Matthew and Shanyce L. Campbell. 2016. Evaluating teacher preparation using graduates' observational ratings. *Educational Evaluation and Policy Analysis* 36(4): 603-625.

Roza, Marguerite and Raegen Miller. 2009. Separation of degrees: State-by-state analysis of teacher compensation for master's degrees. Unpublished paper, Center for American Progress and Center on Reinventing Public Education.

Steinberg, Matthew P. and Rachel Garrett. 2016. Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis* 38(2): 293-317.

Toch, Thomas and Robert Rothman. 2008. Rush to judgment: Teacher evaluation in public education. Unpublished paper, Education Sector Reports.

Weisberg, Daniel, Susan Sexton, Jennifer Mulhern, David Keeling, Joan Schunck, Ann Palcisco, and Kelli Morgan. 2009. The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Unpublished paper, The New Teacher Project.

Whitehurst, Grover J., Matthew M. Chingos, and Katharine M. Lindquist. 2014. Evaluating teachers with classroom observations: Lessons learned in four districts. Unpublished paper, The Brookings Institute.

Table 1: Descriptive Statistics for the Analytic Sample (2005-06 Through 2013-14)

Characteristics	All Teachers	Teachers with Graduate Degrees	Teachers without Graduate Degrees
<i>Teacher-Year Characteristics</i>			
Female	79.87	81.12	79.27
Minority	17.79	18.28	17.56
Age	41.64	44.74	40.16
Teaching Experience	12.38	15.40	10.93
Graduate Degree	32.25	100.00	0.00
NBC	11.11	18.98	7.36
<i>School Characteristics</i>			
School Size	790.06	815.83	777.79
Elementary School	46.58	45.24	47.22
Middle School	19.96	19.75	20.06
High School	29.03	30.51	28.32
Other School Level	4.43	4.51	4.40
Percent Economically Disadvantaged	54.10	52.49	54.87
Percent Minority	47.87	47.63	47.97
Performance Composite	67.02	67.51	66.79
Total Per-Pupil Expenditures	8683.50	8781.80	8636.70
Short-Term Suspension Rates (Per 100 Students)	20.50	19.88	20.80
Teacher-Year Observations	875,092	282,225	592,867

Note: This table displays teacher and school characteristics for the full analysis sample, combined, and for teachers with and without a graduate degree. Teacher characteristics identify unique teacher-year observations; school characteristics identify unique teacher-school-year observations.

Table 2: The Signaling Effect of Graduate Degrees on Teacher Value-Added

	Elem Math	Elem Read	5th Grade Science	Middle Math	Middle Read	8th Grade Science	High Math	High Science	High English	High Social Studies
	<i>Overall Analyses</i>									
Graduate Degree	-0.003 (0.002)	-0.004** (0.001)	-0.002 (0.004)	0.001 (0.002)	-0.005** (0.001)	-0.004 (0.006)	-0.010* (0.004)	-0.013* (0.005)	0.000 (0.002)	-0.013** (0.005)
	<i>In-Area Analyses</i>									
In-Area	0.003 (0.002)	-0.003* (0.001)	0.001 (0.004)	0.012** (0.003)	-0.002 (0.002)	0.013 (0.007)	0.013** (0.005)	-0.001 (0.005)	0.004 (0.003)	-0.000 (0.005)
Out-Area	-0.012** (0.003)	-0.007** (0.002)	-0.018** (0.005)	-0.008** (0.003)	-0.008** (0.002)	-0.010 (0.008)	-0.033** (0.006)	0.004 (0.008)	-0.001 (0.004)	-0.021* (0.009)
Unclassified	-0.013** (0.005)	-0.005 (0.003)	0.009 (0.009)	-0.012* (0.005)	-0.003 (0.003)	-0.017 (0.010)	-0.025** (0.009)	-0.046** (0.009)	-0.010 (0.005)	-0.048** (0.010)
	<i>NBC Teacher</i>									
NBC	0.030** (0.003)	0.007** (0.002)	0.032** (0.005)	0.037** (0.003)	0.015** (0.002)	0.043** (0.008)	0.059** (0.005)	0.058** (0.007)	0.015** (0.003)	0.040** (0.007)
	<i>Standard Deviation of Teacher Value-Added</i>									
	0.161	0.103	0.138	0.144	0.060	0.113	0.159	0.179	0.055	0.172
Observations	2,886,044	3,729,430	912,639	2,908,171	3,114,604	723,338	1,676,671	1,212,861	1,004,175	1,108,138

Note: This table reports differences in adjusted-average student achievement between teachers with a graduate degree and those who have undergraduate degrees only (for NBC teachers, results are in comparison to non-NBC teachers). To help compare value-added and evaluation rating results, this table also reports the standard deviation of individual teacher value-added. * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.

Table 3: The Signaling Effect of Graduate Degrees on Teacher Evaluation Ratings

	Leadership	Classroom Environment	Content Knowledge	Facilitating Student Learning	Reflecting on Practice
<i>Overall Analyses</i>					
Graduate Degree	0.063** (0.004)	0.047** (0.005)	0.060** (0.005)	0.055** (0.004)	0.071** (0.005)
Observations	256,268	150,238	149,607	256,249	149,515
<i>In-Area Analyses</i>					
In-Area	0.073** (0.009)	0.036** (0.011)	0.098** (0.012)	0.078** (0.008)	0.075** (0.012)
Out-Area	0.019 (0.011)	0.058** (0.015)	-0.016 (0.015)	0.004 (0.011)	0.027 (0.016)
Unclassified	-0.048** (0.015)	-0.033 (0.018)	-0.014 (0.019)	-0.021 (0.014)	-0.022 (0.019)
Observations	73,078	43,999	43,847	73,069	43,804
<i>NBC Teacher</i>					
NBC	0.390** (0.005)	0.298** (0.008)	0.400** (0.009)	0.361** (0.005)	0.421** (0.009)
Observations	256,268	150,238	149,607	256,249	149,515

Note: This table reports differences in standardized teacher evaluation ratings between teachers with a graduate degree and those who have undergraduate degrees only (for NBC teachers, results are in comparison to non-NBC teachers). * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.

Table 4: The Human Capital Effect of Graduate Degrees on Teacher Value-Added

	Elem Math	Elem Read	5th Grade Science	Middle Math	Middle Read	8th Grade Science	High Math	High Science	High English	High Social Studies
	<i>Overall Analyses</i>									
Graduate Degree	0.003 (0.004)	0.003 (0.003)	0.021* (0.011)	0.005 (0.004)	0.003 (0.003)	0.004 (0.013)	0.006 (0.009)	0.020 (0.011)	0.013* (0.006)	0.001 (0.012)
	<i>In-Area Analyses</i>									
In-Area	-0.006 (0.006)	-0.000 (0.004)	0.034* (0.015)	0.020** (0.006)	0.004 (0.004)	0.014 (0.021)	0.019 (0.012)	0.038* (0.015)	0.013 (0.009)	0.031* (0.016)
Out-Area	0.008 (0.006)	0.009 (0.005)	-0.001 (0.014)	-0.006 (0.006)	0.000 (0.004)	0.029 (0.016)	-0.009 (0.013)	0.017 (0.016)	0.003 (0.009)	-0.029 (0.018)
Unclassified	-0.003 (0.010)	0.003 (0.007)	0.049 (0.027)	0.006 (0.008)	0.015 (0.007)	-0.100** (0.024)	0.032 (0.020)	-0.033 (0.019)	0.019 (0.015)	-0.032 (0.025)
	<i>NBC Teacher</i>									
NBC	-0.001 (0.005)	0.001 (0.004)	-0.006 (0.015)	-0.001 (0.006)	0.005 (0.004)	-0.012 (0.014)	-0.006 (0.009)	-0.027** (0.009)	-0.002 (0.006)	0.003 (0.010)
	<i>Standard Deviation of Teacher Value-Added</i>									
	0.161	0.103	0.138	0.144	0.060	0.113	0.159	0.179	0.055	0.172
Observations	2,886,044	3,729,430	912,639	2,908,171	3,114,604	723,338	1,676,671	1,212,861	1,004,175	1,108,138

Note: Comparing within teachers, this table reports differences in adjusted-average student achievement before and after teachers earn a graduate degree (for NBC teachers, results compare within teachers before and after earning NBC). To help compare value-added and evaluation rating results, this table also reports the standard deviation of individual teacher value-added. * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.

Table 5: The Human Capital Effect of Graduate Degrees on Teacher Value-Added (Timing Analyses)

	Elem Math	Elem Read	5 th Grade Science	Middle Math	Middle Read	8 th Grade Science	High Math	High Science	High English	High Social Studies
<i>Timing Analyses</i>										
Two Years Before Degree Conferral	0.006 (0.006)	0.006 (0.005)	0.013 (0.015)	-0.004 (0.006)	-0.008 (0.005)	-0.004 (0.019)	0.000 (0.012)	-0.027 (0.016)	-0.006 (0.008)	0.001 (0.016)
One Year Before Degree Conferral	0.014* (0.006)	0.004 (0.005)	0.000 (0.014)	-0.002 (0.006)	-0.007 (0.004)	-0.002 (0.016)	-0.010 (0.012)	-0.024 (0.014)	-0.004 (0.008)	0.008 (0.017)
Year of Degree Conferral	-0.004 (0.005)	0.003 (0.004)	0.006 (0.012)	0.006 (0.006)	-0.009* (0.004)	-0.000 (0.015)	0.003 (0.010)	-0.001 (0.012)	-0.013 (0.008)	0.007 (0.013)
One Year After Degree Conferral	0.004 (0.005)	0.006 (0.004)	0.011 (0.011)	0.004 (0.004)	-0.010* (0.004)	-0.001 (0.014)	-0.015 (0.010)	0.005 (0.010)	0.001 (0.006)	0.003 (0.011)
Two Years After Degree Conferral	-0.006 (0.005)	-0.001 (0.004)	-0.006 (0.010)	0.008 (0.005)	0.001 (0.004)	0.028* (0.013)	-0.006 (0.009)	0.012 (0.009)	0.014* (0.007)	-0.004 (0.010)
Observations	2,886,044	3,729,430	912,639	2,908,171	3,114,604	723,338	1,676,671	1,212,861	1,004,175	1,108,138

Note: Comparing within teachers, this table reports differences in adjusted-average student achievement in the years immediately before earning a graduate degree, in the year of degree conferral, and in the years immediately after earning a graduate degree. * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.

Table 6: The Human Capital Effect of Graduate Degrees on Teacher Evaluation Ratings

	Leadership	Classroom Environment	Content Knowledge	Facilitating Student Learning	Reflecting on Practice
<i>Overall Analyses</i>					
Graduate Degree	0.041** (0.016)	0.015 (0.026)	0.029 (0.027)	0.014 (0.016)	0.006 (0.027)
Observations	256,268	150,238	149,607	256,249	149,515
<i>In-Area Analyses</i>					
In-Area	0.094* (0.043)	0.010 (0.074)	0.115 (0.078)	0.037 (0.043)	0.036 (0.077)
Out-Area	0.039 (0.046)	-0.017 (0.090)	0.007 (0.096)	-0.008 (0.048)	0.020 (0.094)
Unclassified	0.087 (0.076)	-0.020 (0.116)	0.088 (0.127)	0.024 (0.078)	0.054 (0.124)
Observations	73,367	44,262	44,106	73,360	44,060
<i>NBC Teacher</i>					
NBC	0.057 (0.031)	-0.018 (0.067)	0.004 (0.070)	0.011 (0.030)	0.047 (0.072)
Observations	256,268	150,238	149,607	256,249	149,515

Note: Comparing within teachers, this table reports differences in standardized teacher evaluation ratings before and after teachers earn a graduate degree (for NBC teachers, results compare within teachers before and after earning NBC). * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.

Figure 1: Graduate Degrees Held by NCPS Teachers

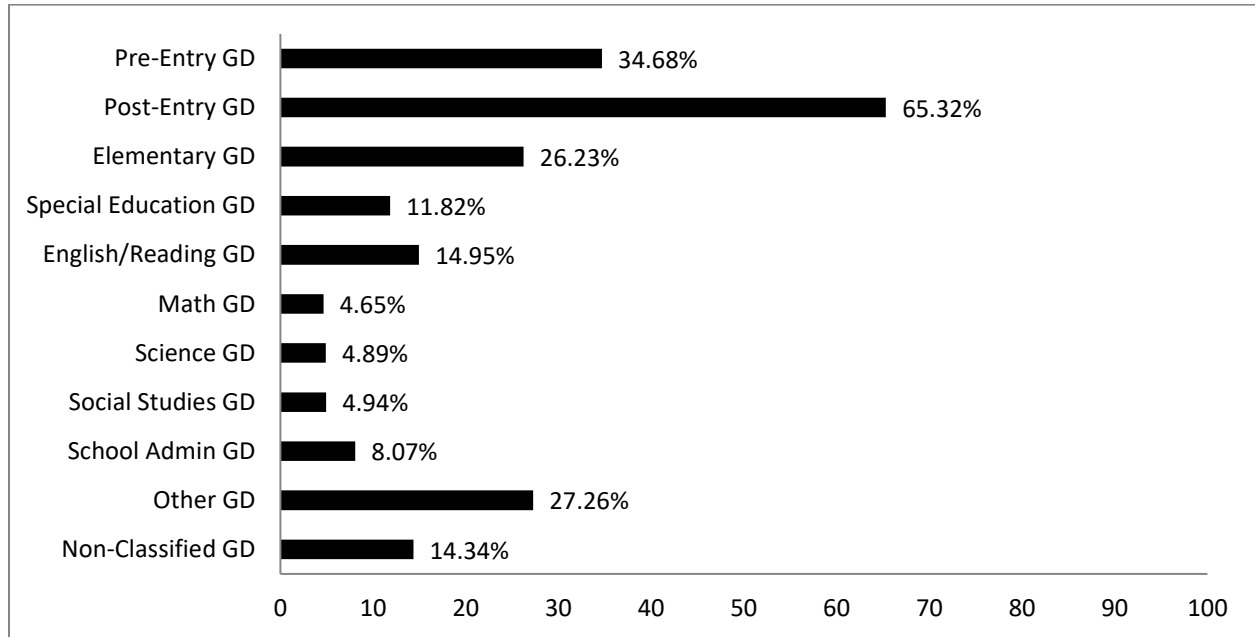
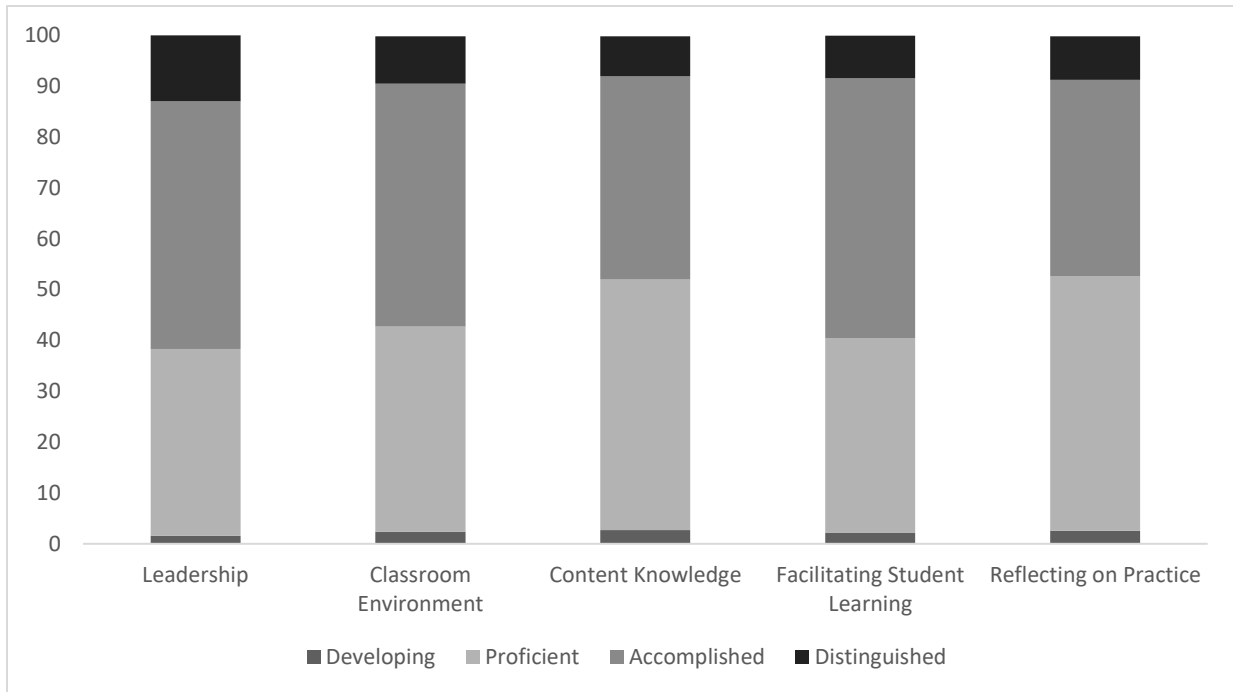


Figure 2: The Distribution of Teacher Evaluation Ratings



Note: For each of the five professional teaching standards, this figure displays the percentage of teachers rating at developing, proficient, accomplished, and distinguished.

Appendix Table A.1: Unique Counts of Teachers Contributing to Model Estimates

	Teacher Value-Added Models										Teacher Evaluation Rating Models	
	Elem Math	Elem Read	5 th Grade Science	Middle Math	Middle Read	8 th Grade Science	High Math	High Science	High English	High Social Studies	Leadership	Classroom Environment
	<i>Signaling Analyses (School Fixed Effect)</i>											
Graduate Degree	26,391	27,720	11,394	13,723	16,273	4,083	7,710	5,361	6,562	4,507	110,332	91,197
In-Area Degree	25,764	27,422	11,101	13,137	15,960	3,376	7,391	5,113	6,319	3,942	37,583	28,853
Out-of-Area Degree	26,003	27,013	11,225	13,678	16,208	4,076	7,699	5,358	6,550	4,504	33,449	25,589
Unclassified Degree	18,342	19,270	7,840	12,606	14,988	3,754	7,496	4,818	6,376	4,378	26,368	20,525
	<i>Human Capital Analyses (Teacher Fixed Effect)</i>											
Graduate Degree	1,972	2,004	1,111	1,163	1,262	336	458	325	390	230	3,672	3,175
In-Area Degree	826	1,135	463	650	709	91	199	149	196	115	631	541
Out-of-Area Degree	1,023	790	586	839	933	238	241	154	183	111	527	430
Unclassified Degree	410	406	228	285	319	81	120	95	101	62	188	164

Note: This table displays the unique number of teachers contributing to the signaling and human capital analyses. I only display counts for the Leadership and Classroom Environment standards since counts for Leadership are very similar to counts for Facilitating Student Learning and counts for Classroom Environment are very similar to counts for Content Knowledge and Reflecting on Practice.

Appendix Table A.2: Signaling and Human Capital Effects of Graduate Degrees on Teacher Value-Added (Content Area)

Graduate Degree	Elem Math	Elem Read	5 th Grade Science	Middle Math	Middle Read	8 th Grade Science	High Math	High Science	High English	High Social Studies
<i>Content-Area Analyses: Signaling Effect</i>										
Elementary	0.005* (0.002)	-0.002 (0.002)	0.001 (0.004)	0.018** (0.004)	0.003 (0.002)	-0.002 (0.014)	-0.015 (0.014)	0.036 (0.032)	-0.005 (0.010)	-0.022 (0.036)
Special Education	-0.020** (0.006)	-0.003 (0.004)	-0.006 (0.014)	-0.008 (0.009)	-0.010* (0.005)	-0.116** (0.036)	-0.027 (0.018)	0.068 (0.051)	-0.018 (0.011)	0.023 (0.024)
Mathematics	-0.022 (0.013)	-0.005 (0.009)	-0.011 (0.022)	0.001 (0.003)	-0.012 (0.008)	0.001 (0.018)	0.012* (0.005)	0.099** (0.022)	-0.040 (0.025)	0.051 (0.091)
Science	-0.006 (0.010)	-0.014 (0.007)	-0.002 (0.017)	-0.009 (0.006)	0.001 (0.007)	0.014* (0.006)	-0.022 (0.017)	-0.003 (0.005)	0.025 (0.018)	-0.049 (0.029)
English-Reading	0.007 (0.004)	-0.002 (0.003)	-0.004 (0.008)	0.024** (0.007)	-0.006** (0.002)	0.007 (0.018)	0.035 (0.024)	-0.097** (0.035)	0.003 (0.003)	-0.029 (0.026)
Social Studies	-0.032** (0.010)	-0.011 (0.007)	-0.015 (0.022)	0.005 (0.009)	-0.006 (0.004)	-0.033 (0.015)	-0.030 (0.033)	-0.103* (0.046)	0.031 (0.014)	0.001 (0.005)
Administration	-0.008 (0.005)	-0.002 (0.004)	-0.032** (0.010)	-0.018** (0.006)	-0.016** (0.004)	0.001 (0.014)	-0.026* (0.010)	-0.049** (0.013)	0.000 (0.007)	-0.016 (0.014)
Other	-0.016** (0.004)	-0.009** (0.003)	-0.006 (0.008)	-0.011* (0.004)	-0.000 (0.002)	-0.002 (0.012)	-0.023** (0.009)	0.005 (0.010)	-0.003 (0.005)	-0.014 (0.014)
Unclassified	-0.012* (0.005)	-0.004 (0.003)	0.009 (0.009)	-0.012* (0.005)	-0.003 (0.003)	-0.016 (0.010)	-0.025** (0.009)	-0.048** (0.009)	-0.011 (0.005)	-0.048** (0.010)
<i>Content-Area Analyses: Human Capital Effect</i>										
Elementary	-0.005 (0.006)	-0.002 (0.005)	0.030 (0.015)	-0.014 (0.019)	-0.001 (0.013)	-0.066 (0.041)	---	---	0.007 (0.051)	0.042 (0.041)
Special Education	0.013 (0.028)	0.004 (0.021)	-0.059 (0.071)	0.013 (0.025)	0.010 (0.020)	-0.064 (0.069)	-0.062 (0.072)	-0.096 (0.070)	-0.027 (0.093)	-0.081 (0.075)
Mathematics	0.005 (0.030)	-0.002 (0.027)	0.057 (0.054)	0.014* (0.007)	-0.003 (0.072)	-0.041 (0.042)	0.017 (0.011)	-0.247** (0.017)	---	---
Science	-0.102 (0.053)	-0.032 (0.029)	0.269* (0.126)	0.014 (0.015)	0.013 (0.029)	0.013 (0.022)	0.017 (0.067)	0.038* (0.015)	0.144** (0.046)	-0.063 (0.112)
English-Reading	0.039** (0.011)	0.004 (0.008)	0.026 (0.026)	-0.000 (0.028)	-0.001 (0.005)	0.019 (0.065)	0.162 (0.104)	0.153** (0.052)	0.011 (0.009)	-0.025 (0.072)
Social Studies	-0.075 (0.089)	0.044 (0.046)	0.233* (0.097)	0.009 (0.031)	-0.011 (0.014)	0.037 (0.037)	-0.074 (0.059)	---	-0.094* (0.037)	0.031* (0.016)
Administration	0.002 (0.009)	0.002 (0.008)	-0.013 (0.022)	0.005 (0.008)	0.005 (0.007)	0.012 (0.023)	-0.020 (0.017)	0.014 (0.020)	-0.012 (0.015)	-0.023 (0.021)
Other	-0.006 (0.010)	0.011 (0.008)	-0.022 (0.023)	-0.005 (0.011)	0.004 (0.006)	0.083** (0.031)	0.019 (0.020)	0.003 (0.023)	0.010 (0.011)	-0.023 (0.030)
Unclassified	-0.004 (0.010)	0.002 (0.008)	0.047 (0.027)	0.007 (0.008)	0.015 (0.007)	0.099** (0.024)	0.031 (0.020)	-0.034 (0.019)	0.017 (0.015)	-0.031 (0.025)
Observations	2,886,044	3,729,430	912,639	2,908,171	3,114,604	723,338	1,676,671	1,212,861	1,004,175	1,108,138

Note: This table reports differences in adjusted-average student achievement (1) between teachers with a graduate degree and those who have undergraduate degrees only (top panel) and (2) before and after teachers earn a graduate degree (bottom panel). * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively. Cells shaded in gray are in-area; in middle grades mathematics and reading, elementary graduate degrees are in-area for 6th grade only.

Appendix Table A.3: Signaling and Human Capital Effects of Graduate Degrees on Teacher Evaluation Ratings (Content Area)

Graduate Degree	Leadership	Classroom Environment	Content Knowledge	Facilitating Student Learning	Reflecting on Practice
<i>Content Area Analyses: Signaling Effect</i>					
Elementary	0.052** (0.006)	0.025** (0.008)	0.045** (0.008)	0.063** (0.006)	0.069** (0.009)
Special Education	-0.045** (0.009)	0.173** (0.012)	-0.173** (0.012)	-0.049** (0.009)	-0.014 (0.013)
Mathematics	0.086** (0.014)	-0.026 (0.019)	0.056** (0.018)	0.067** (0.014)	0.103** (0.019)
Science	0.075** (0.014)	-0.017 (0.018)	0.151** (0.019)	0.084** (0.014)	0.115** (0.018)
English-Reading	0.141** (0.008)	0.111** (0.011)	0.168** (0.011)	0.143** (0.007)	0.176** (0.011)
Social Studies	0.039** (0.013)	0.039* (0.017)	0.108** (0.018)	0.009 (0.013)	0.055** (0.017)
Administration	0.097** (0.010)	0.063** (0.014)	0.069** (0.014)	0.041** (0.010)	0.099** (0.015)
Other	0.050** (0.006)	0.029** (0.008)	0.068** (0.008)	0.043** (0.006)	0.033** (0.008)
Unclassified	-0.028** (0.008)	-0.026** (0.009)	-0.009 (0.009)	-0.017* (0.007)	-0.006 (0.009)
<i>Content Area Analyses: Human Capital Effect</i>					
Elementary	0.036 (0.034)	0.055 (0.055)	0.065 (0.059)	0.036 (0.035)	0.035 (0.063)
Special Education	0.125* (0.057)	-0.007 (0.085)	0.018 (0.080)	0.001 (0.058)	-0.031 (0.082)
Mathematics	0.040 (0.072)	0.045 (0.110)	-0.044 (0.120)	-0.018 (0.076)	-0.047 (0.123)
Science	0.115 (0.093)	0.051 (0.149)	0.036 (0.128)	-0.023 (0.088)	-0.073 (0.150)
English-Reading	0.074 (0.046)	0.103 (0.077)	0.059 (0.074)	0.040 (0.044)	0.120 (0.076)
Social Studies	-0.067 (0.099)	-0.062 (0.127)	0.033 (0.119)	-0.073 (0.093)	-0.000 (0.132)
Administration	-0.076* (0.033)	0.021 (0.062)	-0.033 (0.074)	-0.026 (0.033)	-0.085 (0.072)
Other	0.070* (0.028)	0.002 (0.053)	0.017 (0.056)	0.022 (0.030)	-0.002 (0.051)
Unclassified	0.026 (0.034)	-0.017 (0.053)	0.017 (0.053)	-0.006 (0.034)	0.026 (0.053)
Observations	256,268	150,238	149,607	256,249	149,515

Note: This table reports differences in standardized teacher evaluation ratings (1) between teachers with a graduate degree and those who have undergraduate degrees only (top panel) and (2) before and after teachers earn a graduate degree (bottom panel). * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.

Appendix Table A.4: Signaling and Human Capital Effects of Graduate Degrees on Teacher Evaluation Ratings (Value-Added Sample)

Graduate Degree	Leadership	Classroom Environment	Content Knowledge	Facilitating Student Learning	Reflecting on Practice
<i>Signaling Analyses on Value-Added Sample</i>					
Graduate Degree	0.040** (0.007)	0.031** (0.009)	0.050** (0.009)	0.042** (0.007)	0.047** (0.010)
In-Area	0.073** (0.009)	0.036** (0.011)	0.098** (0.012)	0.078** (0.008)	0.075** (0.012)
Out-Area	0.019 (0.011)	0.058** (0.015)	-0.016 (0.015)	0.004 (0.011)	0.027 (0.016)
Unclassified	-0.048** (0.015)	-0.033 (0.018)	-0.014 (0.019)	-0.021 (0.014)	-0.022 (0.019)
<i>Signaling Analyses Controlling for Individual Teacher Value-Added</i>					
Graduate Degree	0.043** (0.007)	0.035** (0.009)	0.055** (0.009)	0.046** (0.007)	0.051** (0.009)
In-Area	0.073** (0.008)	0.036** (0.011)	0.099** (0.011)	0.078** (0.008)	0.076** (0.012)
Out-Area	0.024* (0.011)	0.064** (0.015)	-0.008 (0.015)	0.010 (0.011)	0.034* (0.016)
Unclassified	-0.038* (0.015)	-0.022 (0.018)	-0.003 (0.018)	-0.010 (0.014)	-0.010 (0.018)
Standardized Value-Added	0.133** (0.003)	0.124** (0.004)	0.136** (0.004)	0.147** (0.003)	0.129** (0.004)
<i>Human Capital Analyses on Value-Added Sample</i>					
Graduate Degree	0.070* (0.033)	-0.019 (0.058)	0.061 (0.062)	0.021 (0.034)	0.020 (0.061)
In-Area	0.080 (0.044)	-0.011 (0.076)	0.104 (0.080)	0.030 (0.044)	0.029 (0.079)
Out-Area	0.052 (0.047)	-0.012 (0.091)	-0.018 (0.096)	-0.001 (0.048)	0.011 (0.094)
Unclassified	0.096 (0.078)	-0.024 (0.113)	0.118 (0.124)	0.039 (0.077)	0.065 (0.122)
Observations	73,078	43,999	43,847	73,069	43,804

Note: For teachers with a value-added estimate, the top panel in this table reports differences in standardized teacher evaluation ratings between teachers with a graduate degree and those who have an undergraduate degree only. The middle panel maintains this sample and controls for a standardized measure of teacher-year value-added. Finally, comparing within teachers, the bottom panel of this table reports differences in standardized teacher evaluation ratings before and after teachers earn a graduate degree. * and ** indicate statistical significance at the 0.05 and 0.01 levels, respectively.