TEACHER PREPARATION POLICIES AND THEIR EFFECTS ON STUDENT ACHIEVEMENT

Gary T. Henry
(corresponding author)
Peabody College
Vanderbilt University
Nashville, TN 37203-5721
gary.henry@vanderbilt.edu

Kevin C. Bastian
Department of Public Policy
    and Education Policy
    Initiative at Carolina
University of North Carolina
    at Chapel Hill
Chapel Hill, NC 27599-2425
kbastian@email.unc.edu

C. Kevin Fortner
Educational Policy Studies
Georgia State University
Atlanta, GA 30302-3982
cfortner2@gsu.edu

David C. Kershaw
Department of Political Science
Slippery Rock University
Slippery Rock, PA 16057
david.kershaw@sru.edu

Kelly M. Purtell
Population Research Center
    and Department of Human
    Development and Family
    Sciences
University of Texas at Austin
Austin, TX 78712
kpurtell@prc.utexas.edu

Charles L. Thompson
Education Policy
    Initiative at Carolina
University of North Carolina
    at Chapel Hill
Chapel Hill, NC 27599-9000
cthomps@email.unc.edu

Rebecca A. Zulli
College of Education
North Carolina State University
Raleigh, NC 27606
rzlowe@ncsu.edu

Abstract
State policies affect the qualifications of beginning teachers in numerous ways, including regulating entry requirements, providing incentives for graduate degrees, and subsidizing preparation programs at public universities. In this paper we assess how these policy choices affect student achievement, specifically comparing traditionally prepared with alternative-entry teachers; in-state traditionally prepared with out-of-state traditionally prepared teachers; teachers beginning with undergraduate degrees with those beginning with graduate degrees; and teachers prepared at in-state public universities with those prepared at in-state private universities. Using school fixed effects to analyze data from North Carolina, we find that: Teach For America corps members are more effective than traditionally prepared teachers; other alternative-entry teachers are less effective than traditionally prepared instructors in high school mathematics and science courses; and out-of-state traditionally prepared teachers are less effective than in-state traditionally prepared teachers, especially in elementary subjects where they constitute nearly 40 percent of the workforce.

1. INTRODUCTION

States establish many policies affecting the preparation of teachers entering the workforce, including regulating entry into the profession, providing incentives to obtain graduate degrees, and subsidizing teacher preparation programs at in-state public institutions of higher education. As recently as twenty years ago, most teachers were prepared through traditional university-based preparation programs, regulated by the state where they would teach. In traditional preparation programs, individuals qualify for a license to teach in the state and concurrently receive a college degree, usually an education major. Now, most states have relaxed the qualifications they require for individuals to begin teaching in public schools and have established procedures for teachers who were prepared at traditional programs in other states to receive reciprocal licenses (Boyd et al. 2006).

These policies have had notable effects on the composition of the teacher workforce. For instance, in 1998–99 the number of new teachers who had entered through alternative routes stood at 10,000; by 2005–06 that number had increased roughly fivefold; and by 2009–10, 39 percent of the teachers who had entered the profession in the last five years had done so through alternative routes (Feistritzer 2011). These changes are especially prominent in North Carolina, a state with a variety of routes into teaching and long-established, nontraditional pathways. Since 1985–86, North Carolina has had a policy to "encourage lateral entry into the profession of teaching by skilled individuals from the private sector" (NCGA 1985, Article 20). This policy eased the way for Eastern North Carolina to become an original Teach For America (TFA) placement site in 1990. As of 2009–10, approximately 15 percent of the state's public school teachers had entered the profession through an alternative route and nearly 29 percent were prepared in traditional programs outside of the state of North Carolina (authors' analysis). A plurality of teachers in North Carolina (48 percent) were prepared through the state's teacher preparation programs, with approximately 75 percent of these teachers coming from in-state public universities (authors' analysis). Clearly, state policies have resulted in substantial variations in the education and type of preparation that teachers have received prior to beginning to teach.

One justification for states to deregulate the teaching profession by reducing barriers to entry has been economic—reducing the costs of entry can increase the labor supply, increase competition in the labor market and, consequently, reduce both teachers' compensation and the expenditures for public schools (Sass 2011). There can be additional economic and social consequences of these state policies, however. On the one hand, removing barriers into teaching may increase the labor supply, reduce pressures on public coffers to increase compensation for public school teachers, increase diversity in the

teaching workforce, and/or improve student achievement if those previously barred from the profession turn out to be effective teachers. On the other hand, removing barriers may reduce the pressure to innovate in the delivery of instruction by keeping the costs of the dominant model of assigning teachers to classes relatively low, increase the churn of teachers through the public schools since lower entry costs mean lower exit costs, and/or reduce student achievement if those previously barred from the profession are less effective teachers. In addition to setting these regulations for entry into the profession, states also influence the teacher workforce and potentially, student achievement, by granting reciprocal licenses for teachers prepared out of state, providing financial incentives for obtaining a graduate degree, and subsidizing the costs of traditional teacher preparation at in-state public institutions of higher education.

Outside of a few studies that we will review in a later section, little is known about the average effectiveness of individuals with varying types of qualifications to teach in public school classrooms. According to a recent National Research Council (NRC) report, the question of the effectiveness of teachers with different qualifications at entry into the teaching profession remains open (Kane, Rockoff, and Staiger 2008; NRC 2010). The limited, available evidence suggests that there is more variation in student test-score gains within the categories typically used to classify teachers' preparation than between these categories (NRC 2010). This has been interpreted to mean that factors other than teacher preparation have a greater influence on teachers' effectiveness. This does not rule out that meaningful differences exist in the effectiveness of teachers with different types of qualifications, however. The numerous routes into teaching, including 130 alternative pathways documented in the NRC review, have been frequently lumped into two types: (1) regular (also known as traditional) certification and (2) alternative certification or lateral entry (NRC 2010). In this study, we use more fine-grained categories that have greater relevance for state teacher policies.

In addition to their breadth, the categories used in prior research often measure teachers' current certification status rather than their preparation prior to entering the classroom, and are thus more fluid than is apparent at first glance. For example, in most states teachers who enter through alternative routes are reclassified as fully certified teachers if they remain in teaching for three or four years, complete the required coursework, and pass exams in content knowledge and/or instructional methods. Two teachers with very different preparation experiences at entry into teaching can be classified into the same category based on their experience and the amount of coursework completed after entering teaching. Therefore, cross-sectional studies that rely on certification status at a given point in time, rather than teachers' initial

preparation, may mask important differences in effectiveness between preparation categories.

Overall, understanding how teacher effectiveness varies by these initial preparation experiences provides important evidence that can influence how teacher licensure regulations are set, how principals and school administrators make hiring decisions, how states allocate funding to teacher education programs, and how both traditional and alternative teacher preparation programs are designed. North Carolina, in particular, has a clear history of being a leader in teacher reform initiatives and teacher quality research. In addition to the early recognition of alternate preparation routes into teaching, including TFA, North Carolina has long encouraged graduate degrees as well as the National Board Certification (NBC) program by offering financial incentives equal to 10 and 12 percent of teachers' base salaries, respectively.[1] North Carolina was also an early proponent of school accountability policies based on student achievement, establishing an accountability system in 1996 with requirements for school ratings, awards, and strong repercussions for failure to meet standards (Carnoy and Loeb 2002; Dee and Jacob 2009).

Importantly, North Carolina's engagement in reforming teacher policies continues today. For example, North Carolina requires teacher preparation programs within the state to continually "upgrade their standards" and report on the effectiveness of their graduates in raising student test scores (Henry et al. 2012). The North Carolina Legislative Education Oversight Commission exercises this responsibility annually by reviewing the state's teacher preparation programs, including the value-added scores of their graduates (Henry et al. 2011). In addition, The Excellent Public Schools Act (NCGA 2012), a sweeping education reform proposal that was partially adopted by the North Carolina General Assembly in 2012 with the remainder under consideration in the 2013 session, proposes significant changes that can be expected to affect the teacher labor force. For example, the act created and funds the North Carolina Teacher Corps—a new program modeled on TFA that recruits recent in-state college graduates to enter teaching—and simplified dismissal of teachers and administrators in low-performing schools. Other portions of this act, currently under consideration in the legislature, include the elimination of teacher tenure and requiring all school districts to adopt pay-for-performance programs. To fund these education reform proposals, some legislators have advocated for a revenue neutral approach that eliminates subsidies for master's degrees and NBC (Luebke 2011). Finally, North Carolina's successful Race to the Top proposal funds an expansion of TFA within the state (Henry et al. 2012).

---

1. Because NBC does not constitute an initial preparation category in that it requires a minimum of three years of teaching experience and this study focuses on teachers in their first three years of experience, we do not include NBC as a teacher preparation category for this study.

In this reform environment the education policy community has a heightened interest in research evidence concerning novice teacher effectiveness. To respond, this study leverages a longitudinal database of teacher licensure, education background, and other data sources to (1) categorize individuals into policy-relevant teacher preparation categories, which include both the formal education and the specific teacher preparation that individuals held *at the time of entry* into the profession, and (2) examine five distinct teacher preparation comparisons that we believe to be of critical importance in understanding the effects of current teacher preparation policy:

1) Are traditionally prepared teachers, which we define as public school teachers who completed the requirements for initial licensure by earning an undergraduate or graduate degree, more or less effective than alternative entry teachers (non-TFA)?
2) Are traditionally prepared teachers more or less effective than TFA corps members?
3) Are in-state traditionally prepared teachers more or less effective than out-of-state traditionally prepared teachers?
4) Are traditionally prepared teachers who begin with graduate degrees more or less effective than those who begin with undergraduate degrees?
5) Are teachers prepared in in-state private institutions of higher education more or less effective than teachers prepared in in-state public institutions of higher education?

Because there is good reason to hypothesize that teachers with different types of preparation may produce better (or worse) results in some grades and subjects than others, we examine these comparisons at the elementary, middle, and high school levels on student mathematics and reading/English tests, as well as high school science and social studies exams. Next, we review selected research on the five specific comparisons we are examining and present our hypotheses for each comparison. Then, we detail the classification scheme that we used to classify teachers into policy relevant preparation categories and focus on the data, sample, and modeling approaches used for the study. Finally, we lay out our findings and present our conclusions.

2. REVIEW OF TEACHER PREPARATION POLICIES AND EXPLANATIONS OF DIFFERENTIAL EFFECTIVENESS

In this section we review state teacher preparation policies that may yield differences in effectiveness, as measured by value-added estimates of teachers' effects on their students' achievement.

Alternate Entry

A primary source of diversification in the qualifications of teachers over the last two decades is the increase in alternatively certified instructors. Except for TFA corps members, who we classify separately, we refer to those individuals who had not completed all requirements for initial licensure prior to entering into the teaching profession as *alternative entry*. These individuals teach and are required to concurrently complete teacher education coursework and pass licensure exams prior to earning full licensure. Due to the rapid increase in the alternatively prepared teacher population, a sizable research body has examined the effectiveness of traditionally versus alternatively certified instructors and returned two main findings: (1) teachers holding regular/traditional certification appear more effective than alternatively certified teachers in the early stages of their careers, but these returns to regular certification fade quickly, and (2) there is more variation in teacher effectiveness within traditionally and alternatively certified categories than between them (Goldhaber and Brewer 2000; Boyd et al. 2006; Clotfelter, Ladd, and Vigdor 2007, 2010; Kane, Rockoff, and Staiger 2008; Constantine et al. 2009).

In their study, which provided the "proof-of-concept" for relating teacher qualifications to student achievement, Goldhaber and Brewer (2000) relied on self-reports of certification status—standard, probationary, emergency, private, and non-certified—by a national sample of teachers to estimate the effects of regular or standard (traditional) versus alternative certification. More recent work in North Carolina elementary and high schools classified teachers as regular, lateral entry, or "other" according to their licensure status (Clotfelter, Ladd, and Vigdor 2007, 2010). This coding scheme improved on past work by including a separate category for teachers who at some point during the study period were classified as lateral entry (i.e., alternative entry) but currently held regular licenses.

Two studies in New York City by Boyd and colleagues and Kane and colleagues advanced the measurement of teacher preparation further. Here, Boyd et al. (2006) classified teachers into six categories (college recommended, individual evaluation, New York City Teaching Fellows, TFA, temporary license, and other) and Kane, Rockoff, and Staiger (2008) included five groups (standard, New York City Teaching Fellows, TFA, international, and uncertified). Importantly, both of these studies treated teacher preparation/certification as a *fixed* trait, anchored to a teacher's initial status upon first being hired in New York State (Boyd et al. 2006) or New York City (Kane, Rockoff, and Staiger 2008). In addition, these papers recognized the current diversity in teacher entry qualifications by including more teacher groups in their classification schemes.

Constantine et al. (2009) used a random assignment procedure to compare alternatively certified teachers to traditionally trained teachers and found no differences between the two groups in comparable elementary classrooms. Many of the "alternatively certified" teachers also had coursework to prepare for work as teachers, although it was less extensive than the traditionally prepared teachers (75–274 hours versus 275–295 hours, respectively). Thus, the difference in preparation between the two types of training is unclear. Finally, Sass (2011) separately examined three groups of early-career, alternatively certified teachers who received different types of alternative training in Florida. Using testing data on students in grades 4–10, he found the performance of some types of alternative certification better and others worse than traditionally certified teachers, depending on the specific program, sample size, subject, type of test, and specification.

It is clear from the existing literature that both the variability in alternative entry programs and the availability of data to separate a teacher's fixed preparation from her time-varying certification status make it challenging to classify teachers and estimate the effects of those categories on student achievement. In this study, we remove some of this variability by excluding TFA corps members from the alternative entry category and capitalize on rich administrative data to classify individuals based on their fixed qualifications upon entry into teaching. Based on literature cited here, which tended to find some small, early-career effectiveness differences between traditionally certified/prepared and alternatively certified/prepared teachers, we hypothesize that early-career alternative-entry instructors will be less effective than early-career traditionally trained teachers.

Teach For America

Perhaps the most studied alternative entry program in the United States, TFA has ambitious goals to increase the number of corps members it annually places—from 7,300 to 13,000—and recently secured $50 million in Investing in Innovation funding to go along with $10 million in matching grants to support this expansion effort (Donaldson and Johnson 2011). Although there is little disagreement that TFA is highly successful at recruiting academically competitive individuals into the teaching profession—many of whom might not have gone into teaching otherwise—and placing them into high-poverty schools, the program engenders sharp critiques from those concerned about insufficient levels of training for corps members (prior to their first year teaching TFA, corps members attend a five-week summer institute and then receive instructional coaching and professional development throughout

their two-year commitment) and the high rates of attrition after their two-year program commitment.

The existing research regarding TFA corps members' effects on student achievement returns mixed evidence (Raymond, Fletcher, and Luque 2001; Darling-Hammond et al. 2005; Boyd et al. 2006; Decker, Mayer, and Glazerman 2006; Kane, Rockoff, and Staiger 2008; Xu, Hannaway, and Taylor 2011). Over time, however, the trend in the evidence suggests that TFA corps members are effective at promoting student achievement growth, especially in STEM courses (mathematics and science) and at the secondary school level (Decker, Mayer, and Glazerman 2006; Xu, Hannaway, and Taylor 2011; Henry, Bastian, and Smith 2012). As evidence concerning the effectiveness of TFA has become more positive, attention has turned to the persistence patterns of their corps members, where evidence suggests that most exit the profession after fulfilling the program's two-year teaching commitment (Donaldson and Johnson 2010; Henry, Bastian, and Smith 2012). Here the concern is that high rates of turnover, particularly in the high-poverty schools where corps members are placed, represent significant losses in terms of resources spent on staff development and teacher recruitment/replacement and has the potential to adversely impact school stability and student achievement (Ronfeldt, Loeb, and Wyckoff 2013).

In this study, we examine the effectiveness of TFA teachers in comparison with those traditionally prepared. Although we do not examine the effects of their short-term teaching commitments, we argue that examining their value added provides important evidence about whether their effectiveness as teachers justifies the churn created in schools caused by their frequent exits.[2] Based on recent evidence, we hypothesize that teachers entering classrooms selected and prepared by TFA will be more effective than traditionally prepared instructors.

Out-of-State Teacher Preparation

One way in which states have mitigated teacher shortages is by importing traditionally prepared teachers from other states (Feistritzer 2011). For example, since 2000–01 the number of out-of-state prepared teachers in North Carolina public schools has increased 36 percent, from 21,316 to 29,066; out-of-state prepared instructors are now the second largest source of teachers in North Carolina public schools. Despite their increased presence in the classroom, little is currently known about the characteristics or effectiveness of out-of-state prepared teachers. Recent work by Sass (2011) indicates that in comparison

---

2. Recent work by Kane, Rockoff, and Staiger (2008) in New York City suggests that the average higher value added of TFA corps members sufficiently compensates for their higher rates of turnover.

with in-state prepared teachers in Florida, out-of-state prepared instructors are significantly less likely to pass the state's certification exams on their first attempt and are significantly less likely to graduate from a most competitive college or university. These findings suggest that out-of-state prepared teachers may perform less well due to the relationship between measures of teacher human capital and teacher value-added estimates (Clotfelter, Ladd, and Vigdor 2007; Goldhaber 2007). In their examination of teacher effectiveness, Goldhaber, Liddle, and Theobald (2013) find that out-of-state prepared teachers are generally no more or less effective than their in-state traditionally trained peers in Washington.

Further investigating the effectiveness of out-of-state prepared instructors, especially in states with teacher shortages, is warranted due to evidence from teacher preparation and labor market research that suggests three explanations for why out-of-state prepared teachers may be less effective than in-state traditionally prepared instructors. First, due to teachers' preferences to work close to home, teacher candidates with less human capital may be forced to relocate to find teaching positions (Boyd et al. 2005; Reininger 2012). Second, recent research suggests that, during teacher preparation, more opportunities to learn the curriculum and engage in teaching practice in settings that match future classroom placement produce more effective novice instructors (Boyd et al. 2009). In comparison with their in-state prepared peers, out-of-state prepared teachers may lack familiarity with the importing state's curriculum, educational environment, and culture, and as a result, produce smaller student test score gains. Finally, out-of-state prepared teachers may be less effective due to teacher turnover—the differential attrition of the most effective out-of-state prepared teachers (presumably returning home to teach) or high rates of turnover and a withdrawal of job-related effort, caused by a lack of commitment to the importing state (Ashenfelter 1978; Boyd et al. 2006). Because of these labor market forces and recent evidence about teacher shortages in North Carolina, we hypothesize that in-state prepared teachers will outperform out-of-state prepared teachers.

Graduate Teacher Preparation Programs

A fourth comparison with direct relevance to state teacher policies is whether fully certified teachers who enter the profession with a graduate degree outperform teachers prepared through undergraduate teacher preparation programs. Comparing the effectiveness of teachers who enter with graduate degrees with the effectiveness of those trained in undergraduate programs is salient for at least four reasons. First, policy makers in every state provide financial incentives to obtain master's degrees by paying teachers who hold graduate degrees

substantially more (Roza and Miller 2009; CSG 2010). For example, a new teacher with a master's degree earned approximately $3,100 more than a new teacher with only a bachelor's degree in the 2007–08 school-year (Council of State Governments 2010). Second, a substantial percentage of teachers across the United States holds the master's degree—in 2003–04, approximately 48 percent of teachers held at least a master's degree (Roza and Miller 2009). This figure was higher in certain states, such as Ohio, New York, and Massachusetts, which require teachers to obtain master's degrees a few years after beginning teaching (Roza and Miller 2009). Third, research shows that approximately 90 percent of advanced degrees held by teachers are graduate degrees in education (CSG 2010). Finally, existing research does not explicitly test whether teachers holding a master's degree upon entry into the profession are immediately more effective than undergraduate prepared teachers.

There is a sizable amount of literature on the value of earning an advanced degree, which shows that students who have teachers with a master's degree generally do not experience larger achievement gains than peers taught by an instructor with an undergraduate degree (Ehrenberg and Brewer 1994; Hanushek 1997; Goldhaber and Brewer 2000; Rivkin, Hanushek, and Kain 2005; Clotfelter, Ladd, and Vigdor 2007; Harris and Sass 2011). Some research indicates a negative association between an advanced degree and student achievement, whereas other work suggests that students who have teachers with a master's degree in mathematics perform better in that subject area (Goldhaber and Brewer 1996; Clotfelter, Ladd, and Vigdor 2007). Importantly, most prior research does not simultaneously distinguish whether an individual's advanced degree came before or after beginning teaching nor does it account for alternative routes into the teaching profession that may impact advanced degree results.

In order to make a more fine-grained comparison, we examine the effectiveness of teachers who enter the teaching profession with a master's degree as compared with teachers who enter the teaching profession without a master's degree (undergraduate). Because prior work has found mixed results for master's degree holders and the content/requirements of undergraduate and graduate teacher preparation are similar, we hypothesize that there will be no differences across the two groups of teachers.

Public versus Private Teacher Education Programs

In addition to regulatory policy and direct incentives, states can influence the qualifications of teachers entering the profession through state and local appropriations that reduce the cost of teacher preparation. In 2008–09, for example, data provided by the Delta Cost Project indicate that the average state

and local appropriation for a full-time equivalency student attending a North Carolina public university was $10,875. In total, 1,643 traditionally prepared (undergraduate only) teachers graduated from the state's public universities in 2009 and began teaching in 2009–10. Assuming these graduates attended their respective universities for their entire collegiate tenure, the total, four-year state and local appropriations for these teachers was over $71 million (Delta Cost Project 2012).[3]

Although there is considerable debate about the philosophical and economic justification for this government (taxpayer) support for the education of teachers, the benefits of higher education are substantial (Poterba 1996; Brunner 1997; Winston 1999; Johnstone 2003; Long 2004; Baum and McPherson 2011; Damon and Glewwe 2011; Schneider and Klor de Alva 2011). For example, Schneider and Klor de Alva find that fully employed young adults in the United States with a college degree earn about 40 percent more, annually, than their counterparts with only some college education, and about two-thirds more than those with only a high school diploma. Although the individual benefits alone may justify government subsidies of higher education costs, higher education is also a public good, generating externalities or broader societal benefits, including greater productivity, higher taxes paid on higher earnings, and a more informed and active citizenry (Brunner 1997; Johnstone 2003; Baum and McPherson 2011; Damon and Glewwe 2011; Schneider and Klor de Alva 2011). For instance, looking solely at the economic costs and benefits of Minnesota's state government investment in higher education, Damon and Glewwe (2011) estimate the state's expenditure of $326 million per year generates benefits ranging between $381 million and $786 million per year.

In the case of teacher education in particular, advocates of government subsidies argue that externalities extend far beyond the financial and intangible societal benefits enumerated here to the schooling of the young, with each teacher contributing to the education of hundreds of students in elementary or secondary schools (Poterba 1996). But this putative benefit becomes a real one only if the teachers educated with public support are effective in the classroom. In the present study, we compare the effectiveness of teachers prepared by North Carolina's public universities with the effectiveness of those prepared at private colleges and universities within the state. Because private institutions are regulated by the state in which they are located and are likely to provide

---

3. In total, there are fifteen institutions within the public university system of North Carolina. To calculate the $71 million figure cited here, we: (1) determined the exact number of traditionally prepared undergraduate teachers graduating from each institution in 2008–09 and beginning teaching in 2009–10; (2) assumed that these teachers were full-time equivalency students over a four-year period (2005–06 through 2008–09); and (3) used higher education cost data provided by the Delta Cost Project to calculate total appropriations per campus and year—expressed in 2009 Consumer Price Index dollars.

similar preparation in terms of courses and student teaching, we hypothesize that in-state publicly prepared and in-state privately prepared teachers will perform similarly.

In the next section, we outline our procedures for classifying teachers into preparation categories.

3. CLASSIFYING TEACHERS FOR TEACHER EFFECTIVENESS COMPARISONS

We classified teachers into the categories discussed in the previous section based on the formal education and specific preparation individuals held when they first began teaching. We first coded whether or not the teacher had met all initial licensure requirements when she began teaching. Next, we coded the type of institution—an in-state public, an in-state private, or an out-of-state university—from which she earned her last degree before first entering the profession. Finally, we coded her highest level of degree—undergraduate or graduate—that she held when first entering the classroom. After coding all early-career teachers using these three distinctions, we created nine, non-mutually exclusive teacher preparation categories (traditional preparation, alternative entry [non-TFA], TFA, in-state prepared, out-of-state prepared, undergraduate-degree prepared, graduate-degree prepared, in-state public-university prepared, and in-state private-university prepared) to test for differences in effectiveness between various combinations of these groups (see table 1 for definitions of each preparation category).

To group teachers into these preparation categories, we relied on administrative data sets from three sources. First, we used institutional data from the University of North Carolina General Administration to identify in-state publicly prepared teachers at the undergraduate and graduate levels. Second, TFA provided us with data identifying their corps members in North Carolina. Finally, we utilized the teacher education, licensure audit, and certified salary files from the North Carolina Department of Public Instruction. From these data sets we used several variables to classify teachers, including the year an individual began teaching; the basis for an individual's original teaching license (had she met all the requirements for initial licensure when she started teaching?); and an individual's graduation year, degree type (undergraduate or graduate), and degree conferring institution type (in-state public, in-state private, or out-of-state university). If an individual earned multiple degrees prior to entering the classroom, we categorized her according to the one closest to beginning teaching. With our focus on five different teacher preparation comparisons, some teachers are classified into more than one preparation category—for example, a teacher in the traditional category is also classified as

Table 1. Teacher Preparation Category Definitions

Category	Definition
Traditional	A North Carolina public school teacher who completed the requirements for initial licensure by earning an undergraduate or graduate degree at an in-state public, in-state private, or out-of-state university.
Alternative Entry	A North Carolina public school teacher who entered the profession prior to completing requirements for initial licensure (Teach For America corps members excluded).
Teach For America	A North Carolina public school teacher who began teaching before completing the requirements for initial licensure and entered the profession through the Teach For America program.
In-State Prepared	A North Carolina public school teacher who completed the requirements for initial licensure at an in-state institution by earning an undergraduate or graduate degree before beginning teaching.
Out-of-State Prepared	A North Carolina public school teacher who completed the requirements for initial licensure at an out-of-state institution by earning an undergraduate or graduate degree before beginning teaching.
Undergraduate Prepared	A North Carolina public school teacher who completed the requirements for initial licensure by earning an undergraduate degree at an in-state public, in-state private, or out-of-state university.
Graduate Degree Prepared	A North Carolina public school teacher who completed the requirements for initial licensure by earning a graduate degree at an in-state public, in-state private, or out-of-state university.
In-State Public University Prepared	A North Carolina public school teacher who completed the requirements for initial licensure by earning an undergraduate or graduate degree at an in-state public university.
In-State Private University Prepared	A North Carolina public school teacher who completed the requirements for initial licensure by earning an undergraduate or graduate degree at an in-state private university.

in-state or out-of-state, undergraduate or graduate, and if applicable, in-state public or in-state private university—whereas other teachers are in only one group (e.g., alternative entry or TFA).

To illustrate the size of these policy-relevant teacher preparation categories, in figure 1 we display unique teacher counts from the most recent year of available data (2009–10). This figure shows that three-fourths of North Carolina's approximately 100,000 public school teachers were traditionally prepared. Of these traditionally prepared teachers, 62 percent were prepared in state and 38 percent were prepared out of state; 88 percent began teaching with an undergraduate degree and 12 percent entered the profession holding a graduate degree. In-state prepared teachers were further subdivided into 73 percent from public universities and 27 percent from private universities. Finally, approximately 15 percent of the public school workforce entered the profession alternatively—many of these teachers coming through various alternative
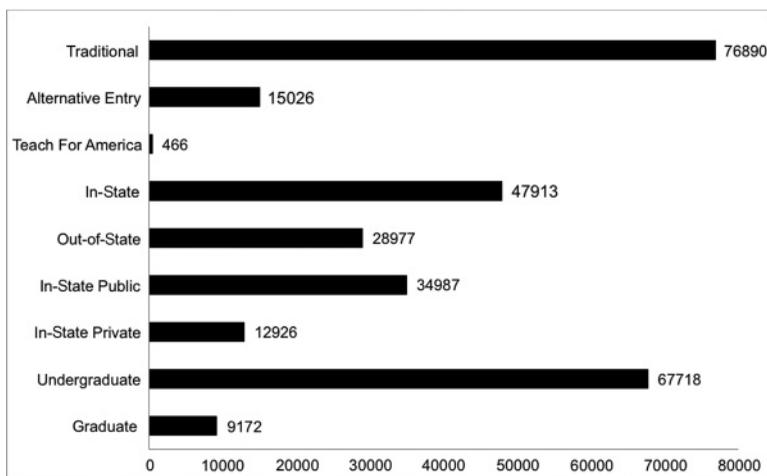
Figure 1. The Distribution of Teacher Categories. *Note*: In the 2009–10 school year, there were 100,633 unique individuals paid as teachers in North Carolina public schools. This figure displays unique teacher counts for each of the teacher categories.

preparation programs and a small number through TFA. Overall, North Carolina has a great diversity in how and where its teachers were prepared to teach.

4. STUDY SAMPLE, DATA, AND METHODS

In this study we estimate the effects of teachers from different preparation categories on the test scores of students in elementary, middle, and high school grades. The central research question is: How do the adjusted-average test score gains of students taught by teachers who have entered teaching through one of the categories (i.e., traditional) compare with the gains of students taught by teachers entering the profession through a relevant comparison category (i.e., alternative entry)? We test for effectiveness differences between the following groups of teachers:

1) Traditionally prepared versus alternative-entry teachers (non-TFA);
2) Traditionally prepared versus TFA corps members;
3) In-state traditionally prepared versus out-of-state traditionally prepared teachers;
4) Undergraduate-degree prepared versus graduate-degree prepared teachers; and
5) In-state public-university prepared versus in-state private-university prepared teachers.

Study Sample

The main objective for this study is to compare the average effect of teachers who entered the profession through different policy relevant preparation categories upon student achievement in eight tested grade-subject combinations: elementary grades mathematics and reading; middle grades mathematics and reading; and high school mathematics (algebra 1, algebra 2, and geometry), science (biology, chemistry, physical science, and physics), social studies (U.S. history and civics/economics), and English. For this we built a statewide analysis file for the 2005–06 through 2009–10 school years in which we: (1) linked students and teachers using actual class rosters, which allowed us to match students to approximately 93 percent of individual instructors over the five-year period; (2) matched students' test scores to their prior test scores, which allowed us to estimate the additional learning or value added during each of the academic years being studied; and (3) constructed numerous other student, teacher, and school variables, which we identify later in the section on covariates, to isolate the effect of our teacher preparation categories.

To further these efforts to identify the effect of teacher preparation, we limit our analyses to teachers with less than three years of experience (first-, second-, or third-year teachers). This decision is based on (1) the expectation that as teachers gain more experience, their initial preparation/qualifications have less influence on their effectiveness and (2) the desire to provide information about teachers proximate to their entry into the profession, since beginning teachers are now the modal category in the teaching workforce (Ingersoll and Merrill 2010). After removing teachers with more than three years of experience from the data set, over 1.7 million test score records for 1.18 million unique students and 22,078 unique teachers were utilized for the analysis.

Table 2 presents descriptive information concerning the levels of prior student performance and the percentage of students qualifying for subsidized school lunches in the classrooms and schools in which our sample of early-career teachers is employed. Here, in comparison to traditionally prepared teachers, alternative entry instructors and TFA corps members work in classrooms and schools with significantly more low-performing and high-poverty students across all grade levels. Out-of-state prepared teachers generally work in environments comparable to in-state prepared teachers—in elementary and middle schools they work in classrooms and schools with slightly lower levels of poverty. In elementary grades, individuals holding graduate degrees upon entry into the profession are employed in classrooms and schools with higher-performing and lower-poverty students; in middle and high school, these differences are mostly at the classroom level. Finally, in comparison to their publicly prepared peers, in-state private university-prepared teachers

Table 2. Teacher Preparation Category Descriptive Information

Teacher Preparation Categories	Elementary School				Middle School				High School			
	Students' Prior EOG Scores	Classroom Poverty Percentage	School Performance Composite	School Poverty Percentage	Students' Prior EOG Scores	Classroom Poverty Percentage	School Performance Composite	School Poverty Percentage	Students' Prior EOG Scores	Classroom Poverty Percentage	School Performance Composite	School Poverty Percentage
Traditional	−0.142 (0.502)	53.95 (27.12)	59.66 (15.83)	57.43 (24.52)	−0.214 (0.729)	47.36 (29.17)	59.78 (15.87)	50.61 (20.53)	−0.047 (0.621)	38.14 (24.92)	68.09 (13.87)	39.09 (19.47)
Alternative Entry	−0.518** (0.716)	56.59** (34.66)	54.86** (16.73)	64.10** (24.78)	−0.505** (0.749)	56.98** (30.31)	51.73** (17.49)	59.28** (21.53)	−0.231** (0.643)	44.04** (26.54)	62.13** (15.71)	45.78** (21.83)
TFA	−0.490** (0.307)	72.53** (22.70)	44.26** (10.45)	85.34** (14.49)	−0.646** (0.659)	69.20** (25.99)	41.07** (11.05)	76.74** (14.25)	−0.519** (0.518)	63.15** (21.08)	50.60** (13.77)	63.94** (23.00)
In-State Prepared	−0.142 (0.491)	54.71 (26.51)	59.75 (15.40)	58.78 (23.68)	−0.220 (0.712)	48.11 (28.85)	59.53 (15.43)	52.03 (19.80)	−0.045 (0.618)	38.10 (24.46)	68.41 (13.59)	38.97 (18.90)
Out-of-State Prepared	−0.142 (0.518)	52.77** (27.98)	59.53 (16.41)	55.49** (25.57)	−0.206 (0.753)	46.34** (29.56)	60.12 (16.47)	48.60** (21.36)	−0.053 (0.628)	38.21 (25.92)	67.41* (14.43)	39.35 (20.67)
Undergraduate Degree Prepared	−0.146 (0.495)	54.29 (26.91)	59.41 (15.70)	58.05 (24.28)	−0.218 (0.717)	47.53 (28.99)	59.86 (15.71)	50.71 (20.41)	−0.068 (0.606)	39.04 (24.73)	67.95 (13.55)	39.80 (19.26)
Graduate Degree Prepared	−0.104** (0.563)	50.56** (28.84)	62.04** (16.80)	51.67** (26.02)	−0.184* (0.819)	46.06* (30.48)	59.10 (17.15)	49.70 (21.44)	0.037** (0.673)	34.40** (25.36)	68.69 (15.12)	36.13** (20.03)
In-State Public University Prepared	−0.161 (0.503)	55.29 (27.15)	59.20 (15.62)	59.18 (24.17)	−0.222 (0.708)	48.46 (28.89)	59.50 (15.38)	52.09 (19.87)	−0.044 (0.613)	38.41 (24.11)	68.53 (13.38)	39.34 (18.73)
In-State Private University Prepared	−0.099** (0.459)	53.33** (24.86)	61.06** (14.80)	57.82* (22.43)	−0.212 (0.729)	46.37** (28.60)	59.65 (15.67)	51.73 (19.46)	−0.043 (0.639)	36.93 (25.75)	67.97 (14.39)	37.56* (19.48)

Notes: For these descriptive statistics the students' prior EOG scores and classroom poverty percentage identify unique teacher and classroom observations; the school performance composite and school poverty percentage identify unique teacher and school observations. Reported significance is in reference to traditional, in-state prepared, undergraduate-degree prepared, and in-state public-university prepared teachers.

*Statistically significant at the 5% level; **statistically significant at the 1% level.

work in classrooms and schools with fewer low-performing and high-poverty students in elementary grades; in middle and high school their work environments are similar. Overall, table 2 indicates that some sorting occurs in the school working environments between individuals with different types of preparation, presumably due to the preferences of administrators and the individual teachers. These observed differences, and the potential for unobserved differences, influence our choice of identification strategy discussed subsequently.

Study Data

Outcome Variables

For this analysis, students' current and prior test score performance is based on the North Carolina grade 3 pre-test, End-of-Grade (EOG) tests, and End-of-Course (EOC) tests. These assessments were developed using the standards and curriculum set by the North Carolina State Board of Education and using psychometric techniques that are commonly used by states to implement federal and state accountability requirements. Elementary (3–5) and middle (6–8) grades models include test scores in mathematics and reading from 2005–06 through 2009–10.[4] A distinct advantage of North Carolina high school data is that criterion-referenced EOC tests can be linked back to specific teachers and their students. High school grades (9–12) analyses included test observations across all ten EOC-tested subjects. Mathematics, social studies, and English test scores are available for all five years, 2005–06 through 2009–10. Science courses were excluded from the analysis in 2006–07 due to test piloting during that year. Finally, to remove secular trends or other year-to-year anomalies in the testing process, we standardized all test scores by year, grade, and subject (elementary and middle grades) or by year and subject (high school grades) and included year fixed effects in model specifications.

Covariates

Recent studies assessing identification strategies have shown that covariate adjusted estimates, when rich covariates are available, substantially reduce bias in effect estimates in comparison with estimates from a randomized control trial (Glazerman, Levy, and Myers 2003; Shadish, Clark, and Steiner 2008; Bifulco 2012). Therefore, we use a rich set of covariates (table 3), coupled with a school fixed effect modeling approach (fully described in the next section), to adjust for many of the factors potentially confounding teacher preparation

---

4. Value-added results are not available for third grade in 2009–10 because North Carolina discontinued the administration of the third-grade EOG pre-test.

Table 3. Covariates Used in the Value-Added Models

Student Covariates	Classroom/Teacher Covariates	School Covariates
1) Prior test scores	1) Number of students	1) School size
2) Classmates' prior test scores	2) Advanced curriculum	2) School size squared
3) Gender	3) Remedial curriculum	3) Violent acts per 1,000 students
4) Race/ethnicity	4) Heterogeneity of prior achievement within the classroom	4) Short-term suspension rate
5) Gifted		5) Total per-pupil expenditures
6) Disability		6) District teacher supplements
7) Currently limited English proficient	5) Single-year indicators for teacher experience	7) Racial/ethnic composition
8) Previously limited English proficient	6) Teaching out-of-field	8) Concentration of poverty
9) Structural mobility	7) Teacher preparation categories	
10) Within year mobility		
11) Between year mobility		
12) Days absent		
13) Overage for grade		
14) Underage for grade		
15) Poverty status		

category estimates. Here, we briefly describe the student, classroom/teacher, and school covariates used to balance the differences in classroom and school environments for teachers from different preparation categories.

At the student level we control for an individual student's prior test score in the subject being analyzed (elementary school), prior scores from both mathematics and reading (middle school), or a student's eighth grade mathematics and reading scores (high school). Using roster information to identify a student's peers, we also control for the average prior performance of all the students in a classroom minus student i. Other student covariates include a continuous measure of days absent and indicators for race/ethnicity (black, Hispanic, Asian, American Indian, and multiracial), gender, subsidized lunch status (free or reduced price), mobility (structural, within-year, between-year), giftedness/disability, limited English proficiency (both currently and previously qualifying), and being underage or overage for grade (skipping a grade or being retained).

To adjust for differences in the individual characteristics of teachers we included our focal teacher preparation category indicators, single-year indicators

of teacher experience (variables for second- and third-year teachers in reference to first-year instructors) and an indicator for teaching out-of-field (not holding a license for the grade level/subject area being taught). For classroom differences we control for the dispersion or range of students' prior test scores within a classroom and the number of students in the classroom. We also include indicators for a classroom's curriculum status (advanced or remedial) in middle and high school analyses.

Finally, to balance differences across the school environments in which teachers from our preparation categories work, we control for the time-varying percentage of students within a school by ethnicity and free/reduced-price lunch status, school size, total per-pupil expenditures, the average local supplement paid to teachers in a school, and two indicators of school orderliness (suspensions per 100 students and reported violent acts per 1,000 students). Because we utilize a school fixed effects approach as our preferred estimation strategy, coefficients for these variables cannot be interpreted as the average, sample-wide differences in student achievement associated with differences in a school covariate but rather how a change in a school characteristic within a particular school is associated with changes in student achievement at the same school.

Methodology

For this study we sought an identification strategy and modeling approach that would yield unbiased or consistent estimates of the effects of our policy-relevant teacher preparation categories on student achievement. A chosen specification should mitigate the confounding effect of variables that influence student performance and may not be balanced across the types of students, classes, and schools in which our sample of teachers from different preparation categories work. To fulfill our modeling objective we considered three specifications: (1) a rich covariate ordinary least squares (OLS) value-added model; (2) a student fixed effect model with classroom/teacher, school, and time-varying student characteristics; and (3) a school fixed effect model with a rich set of student, classroom/teacher, and time-varying school covariates.

A major advantage of the OLS specification is that the sample included in the estimation is the statewide population of early-career teachers. In contrast, school fixed effects models include only those schools in which teachers from the two groups being compared work during the study period, thus potentially excluding many novice teachers from the comparisons. However, research documenting the sorting of teachers across and within districts, the nonrandom assignment of students to teachers, and the influence of school leadership and working conditions on teacher effectiveness and retention suggest that a rich

set of covariates may be insufficient, by themselves, to adjust for variables confounding preparation category estimates (Lankford, Loeb, and Wyckoff 2002; Clotfelter, Ladd, and Vigdor 2005; Rothstein 2010; Boyd et al. 2011; Ladd 2011). Therefore, to better account for unobserved confounders we prefer a school fixed effects specification; we also present rich covariate OLS estimates as a specification check with the more inclusive sample of early-career teachers.

Student fixed effects use each student as his or her own control, thereby eliminating from estimates the influence of unobserved, non-time-varying differences between students—a key benefit given the research evidence concerning the nonrandom assignment of students to schools and classrooms (Wooldridge 2009). An important implication of using student fixed effects for these analyses, however, is that the coefficient on a binary indicator for a teacher preparation category is only identified on the subset of students who have teachers from that preparation category and the relevant comparison category (e.g., in-state and out-of-state prepared teachers). For a specific model (e.g., in-state versus out-of-state prepared teachers) this omits students who have been taught by only one type of teacher and students who have been taught by teachers from different preparation categories than those being compared. Furthermore, difficulties with the student fixed effects are especially amplified by our focus on early-career teachers—in order to estimate the effect of the teacher preparation categories on student test scores, a student must be taught by early-career teachers (fewer than three years' experience) from both of the preparation categories included in the comparison.[5] Therefore, considering both the sample restrictions caused by the student fixed effects approach and the extensive set of student covariates that we already have available (limiting the benefit of the student fixed effect), we ruled out student fixed effects as our primary specification, but include it as a specification check since it removes a potential source of confounding in the restricted sample.

For this study our preferred identification strategy is a value-added model with school fixed effects and a rich set of student, classroom, and time-varying school covariates. This school fixed effect limits comparisons to early-career teachers working in the same schools, thereby adjusting out uncontrolled for, time-invariant school factors, such as teacher sorting, school leadership, or working conditions, that may confound our teacher preparation category

---

5. Teachers with fewer than three years of experience account for approximately 17 percent of the North Carolina teaching workforce during our study period. Based on prior research documenting the sorting of more experienced teachers to higher-performing and lower-poverty students, this suggests that the sample of students matched to early-career teachers in consecutive school years is likely quite different than the statewide population of students (Clotfelter, Ladd, and Vigdor 2005).

estimates. Given the differences in school environments for our preparation categories shown in table 2, we argue that this school fixed effects approach, coupled with our extensive set of covariates, best isolates the effect of our teacher preparation categories while retaining a sample of observations that is sufficiently large and plausibly reflects the diversity of exposure to the types of teachers being compared. For all analyses we use cluster-adjusted standard errors at the school-year level to account for the clustering of students and teachers within schools.

Below, we specify the equation for our school fixed effect value-added model. Prior to doing so, however, we should make clear that our effect estimates represent the adjusted average differences in the effects of teacher preparation categories, regardless of whether those differences are due to selection into or preparation received in the teacher preparation category. We argue that it is this combined effect that is most important to policy makers when they set qualifications for entry into teaching. While it may also be desirable to separate the effects of teachers' preparation from the effects of selection into a preparation category, that is not the goal of the present study. Equation 1 is used to estimate the average effect of our policy-relevant teacher preparation categories:

$$Y_{ijst} = \hat{a}_o + \hat{a}_1 Y_{it-n} + \hat{a}_2 TeacherPrep + \hat{a}_3 X_{ijst} + \hat{a}_4 Z_{jst}$$
$$+ \hat{a}_5 W_{st} + \grave{e}_s + \mathring{a}_{ijst} \tag{1}$$

where Y_{ijst} is the test score for student i in classroom j in school s at time t;

Y_{it-n} represents the prior test scores for student i;

$\hat{a}_2$ estimates the average effect of a teacher preparation category in relation to its reference teacher preparation category (traditional vs. alternative entry and TFA, in-state versus out-of-state, undergraduate versus graduate, and in-state public versus in-state private university);

TeacherPrep is an indicator variable equal to 1 if the teacher entered the profession through that category and 0 if not;[6]

X_{ijst} represents a set of individual student covariates;

Z_{jst} represents a set of classroom and teacher covariates;

W_{st} represents a set of time varying school covariates;

$\grave{e}_s$ is a school fixed effect included to adjust for time-invariant school factors;

and $\mathring{a}_{ijst}$ is a disturbance term representing all unexplained variation.

6. Traditional, in-state, undergraduate, and in-state public university prepared teachers serve as the reference categories for their respective models in comparison to alternative entry and TFA, out-of-state, graduate degree, and in-state private-university prepared teachers.

In addition to estimating the fixed, average effect of a teacher preparation category in reference to its comparison preparation category, we were also interested in examining the variation in teacher effectiveness within and between preparation categories. Although the average effect of each preparation category is the most policy-relevant comparison, especially if the differences are meaningful in magnitude, the degree of overlap in the distribution of teacher effectiveness between teacher preparation categories is also germane. To examine variation we estimated individual teacher effectiveness using an OLS value-added model with a rich set of student, classroom, and school covariates. Here, we specified students' residuals as the measure of individual teacher effectiveness and aggregated these residuals up to the preparation category level to generate parameters of interest—mean, standard deviation, twenty-fifth and seventy-fifth percentiles—for each preparation category within our eight value-added models. Whereas our analyses to estimate the average, fixed effect of each preparation category included a school fixed effect to adjust for unmeasured school factors confounding preparation category estimates, for this work we purposefully excluded a school fixed effect from the specification. This decision is in line with research showing that including a school fixed effect removes substantial variation in individual teacher value-added estimates and compares each teacher's effect to the mean effect within the school rather than the population mean (McCaffrey et al. 2004).

In our findings section we present results for each of our comparisons from our preferred school fixed effects approach, followed by results from rich covariate OLS and student fixed effects specification checks. In all results tables we specify the number of observations used to estimate the teacher preparation coefficients of interest—in rich covariate OLS models, this is simply an observation count, and in fixed effects approaches this is the number of observations experiencing within-unit (school or student) variation for the focal teacher preparation variables. Additionally, we include an appendix table (appendix table A.3) displaying unique teacher counts from each of our value-added model specifications. Finally, we include figures—showing the mean and teacher effectiveness at the twenty-fifth and seventy-fifth percentiles—to illustrate the distribution of teacher effectiveness by preparation category.

5. FINDINGS

Traditional versus Alternative Entry

Consistent with our hypothesis, table 4 shows that teachers entering the profession alternatively are significantly less effective than traditionally prepared teachers in three of eight comparisons—middle grades mathematics, high school mathematics, and high school science.

Table 4. Comparisons of Average Effectiveness for Policy Relevant Preparation Categories (School Fixed Effects)

Policy Relevant Comparison	Elementary Math	Elementary Reading	Middle Math	Middle Reading	High School Math	High School Science	High School English	High School Social Studies
Alternative vs. Traditional	0.008	0.015	−0.016*	−0.001	−0.025*	−0.051**	0.003	−0.010
	(0.010)	(0.008)	(0.007)	(0.004)	(0.012)	(0.017)	(0.007)	(0.013)
Student Observations Used	166,621	237,131	203,774	226,457	190,760	111,384	82,685	110,055
TFA vs. Traditional	**0.073****	**0.038***	**0.136****	0.016	**0.194****	**0.191****	**0.081****	0.086
	(0.023)	(0.017)	(0.023)	(0.014)	(0.034)	(0.046)	(0.024)	(0.049)
Student Observations Used	25,686	34,268	14,623	18,677	13,889	5,727	5,620	8,773
Out-of-State vs. In-State	**−0.019****	**−0.009***	0.000	−0.004	**−0.041****	−0.035	−0.008	−0.030
	(0.005)	(0.004)	(0.009)	(0.005)	(0.014)	(0.027)	(0.010)	(0.016)
Student Observations Used	339,319	463,489	117,642	131,369	107,634	35,178	43,428	89,849
Graduate Degree vs. Undergraduate	−0.003	0.003	**−0.031***	**−0.018***	0.016	**0.059***	0.016	−0.005
	(0.007)	(0.006)	(0.015)	(0.007)	(0.019)	(0.025)	(0.011)	(0.019)
Student Observations Used	200,308	268,174	53,551	82,838	75,267	29,570	35,144	68,557
In-State Private University vs. In-State Public	−0.009	−0.009	−0.012	0.013	−0.009	0.070	0.016	0.023
	(0.007)	(0.006)	(0.015)	(0.010)	(0.021)	(0.048)	(0.011)	(0.019)
Student Observations Used	154,318	219,349	33,247	34,856	54,743	11,787	25,228	40,544

Notes: In these analyses, the second category in each row is the reference category (traditionally-prepared, in-state-prepared, undergraduate-degree prepared, and in-state public-university prepared teachers). *Student Observations Used* indicates the number of students experiencing within-unit (school) variation for the comparison.
*Statistically significant at the 5% level; **statistically significant at the 1% level.

In the remaining five comparisons, traditionally prepared and alternative-entry teachers perform similarly. This suggests that more extensive preparation to teach prior to entry into the profession has a positive impact on student achievement in STEM subjects in secondary grades. The policy and practical significance of these results is increased when we consider the high concentrations of alternative-entry teachers in mathematics and science courses at the secondary level—for instance, alternative-entry instructors are the largest source of early-career teachers in high school science and they taught more than 100,000 students during the study period (see appendix table A.3 for unique teacher counts from the value-added models).

Traditional versus Teach For America

As expected from prior research, our results in table 4 show that TFA corps members are significantly more effective than traditionally prepared teachers in six of eight comparisons—elementary mathematics and reading, middle grades mathematics, and high school mathematics, science, and English—and no different in the remaining two comparisons. Of note here is the magnitude of the TFA effects: Compared with students assigned to a traditionally prepared teacher, students instructed by TFA corps members annually gain approximately 18, 11, and 73 days of additional learning in elementary grades mathematics and reading, and middle grades mathematics, respectively.[7] Overall, although TFA is the smallest source of teachers in North Carolina public schools, accounting for approximately 0.50 percent of the workforce, evidence clearly suggests the program (1) provides highly effective teachers and (2) may provide significant clues for ways to improve teacher selection and preparation (Dobbie 2011).

In-State versus Out-of-State Prepared

Consistent with our hypothesis, results in table 4 show that teachers traditionally prepared out-of-state are significantly less effective than their traditionally prepared in-state peers in three of eight comparisons—elementary mathematics and reading, and high school mathematics. In the remaining five comparisons, in-state and out-of-state prepared teachers performed similarly. These elementary school results are particularly salient due to the concentration of out-of-state prepared teachers in those grades. Overall, out-of-state prepared instructors constitute 36 percent of the early-career elementary school teacher workforce in grades 3–5, and during the study period they taught nearly 115,000 students.

---

7. See Henry et al. (2011) for more information regarding days of equivalent student learning calculations.

Undergraduate versus Graduate-Degree Prepared

Contrary to our hypothesis, results in table 4 indicate that effectiveness differences do exist between teachers entering the profession with traditional teacher preparation at the undergraduate level versus those with preparation at the graduate degree level. In middle grades mathematics and reading those holding a graduate degree are, on average, significantly less effective. Conversely, graduate degree holders are significantly more effective in high school science. Whether these differing results in middle and secondary grades are attributable to differences in the focus of the master's degree—a content-specific or education graduate degree—may represent an opening for future research.

In-State Public University versus In-State Private University Prepared

As expected, table 4 shows that individuals earning teacher preparation degrees at in-state private universities performed similarly to individuals earning teacher preparation degrees at in-state public universities across all eight grade level/subject-area models. In one comparison, high school science, the magnitude of the coefficient was large, suggesting that privately prepared teachers were more effective, but the result was not significant. In relation to our other policy-relevant comparisons these findings indicate that, on average, effectiveness differences are not between in-state prepared teachers from public versus private institutions but rather are between those prepared in-state versus out-of-state or those traditionally prepared versus alternatively prepared.

Specification Checks

In order to assess the potential for bias in these preferred school fixed effect estimates we present results from two types of specification checks in appendix tables A.1 and A.2. First, we specify a rich covariate OLS model to ensure that our school fixed effect approach does not mask effectiveness differences between preparation categories. For example, as reasoned by Goldhaber, Liddle, and Theobald (2013), it is possible that large effectiveness differences exist between preparation categories (e.g., in-state and out-of-state prepared), statewide, but if schools employ teachers of similar effectiveness—the least effective teachers from a high value-added category and the most effective teachers from a low value-added category—the within school comparison may not show effectiveness differences between the preparation groups. Second, we implement a student fixed effect model to eliminate any unobserved, nontime-varying differences between students. Here, following Clotfelter, Ladd, and Vigdor (2007, 2010), we specify a levels model—current, standardized test score as the dependent variable with a student fixed effect (dichotomous variable) and a rich set of classroom, school, and time-varying student covariates—in elementary

and middle grades mathematics and reading and in high school mathematics, science, and social studies.[8] Estimates in elementary and middle grades compare test score outcomes for students with teachers from different preparation categories over time, whereas estimates at the high school level compare test score outcomes across courses (e.g., U.S. history and civics/economics for high school social studies) for students taking multiple EOC exams with teachers from different preparation categories.

To better understand how results may differ across model specifications due to changes in the estimation sample we include unique case counts in each results table and unique teacher counts for each value-added specification in appendix table A.3. For example, when comparing in-state versus out-of-state prepared teachers in elementary grades mathematics, a rich covariate adjustment OLS model uses 403,502 student observations; 4,062 unique in-state prepared and 3,165 unique out-of-state prepared teachers contribute to the focal coefficient. When using fixed effects the student observations experiencing within-unit variation drop to 339,319 and 42,333 observations for school and student fixed effects, respectively. Additionally, the unique teacher counts decrease to 3,148 and 2,201 for in-state prepared teachers and 2,992 and 2,212 for out-of-state prepared teachers for school and student fixed effects, respectively. These changes in the sample, particularly the substantially reduced case counts in the student fixed effect models, increase standard errors and may reduce the generalizability of estimates.

Overall, results from the rich covariate OLS models were consistent with our preferred school fixed effects estimates. Of the fifteen significant findings in table 4, thirteen remain significant in appendix table A.1. When estimates that were not significant in the school fixed effect models were significant in the OLS approach (e.g., alternative entry in high school social studies or TFA in middle grades reading), they were still consistent with the overall direction of findings for the preparation category. Finally, as anticipated in section 4, findings from the robustness checks with student fixed effects were limited due to the substantially reduced sample sizes and increased standard errors of these models. Although only three of the fifteen significant estimates from table 4 were significant with the student fixed effects approach (shown in appendix table A.2), the magnitude and direction of the preparation category results was generally consistent across our fixed effect specifications. For example, the effect of TFA corps members in high school science remains large (0.142

---

8. We cannot estimate a student fixed effects model in high school English because there is only a single North Carolina EOC exam for that subject area (English 1). There are multiple exams in our high school mathematics (algebra 1, algebra 2, geometry), science (biology, chemistry, physical science, physics), and social studies (U.S. history and civics/economics) models, allowing the use of a student fixed effects specification.
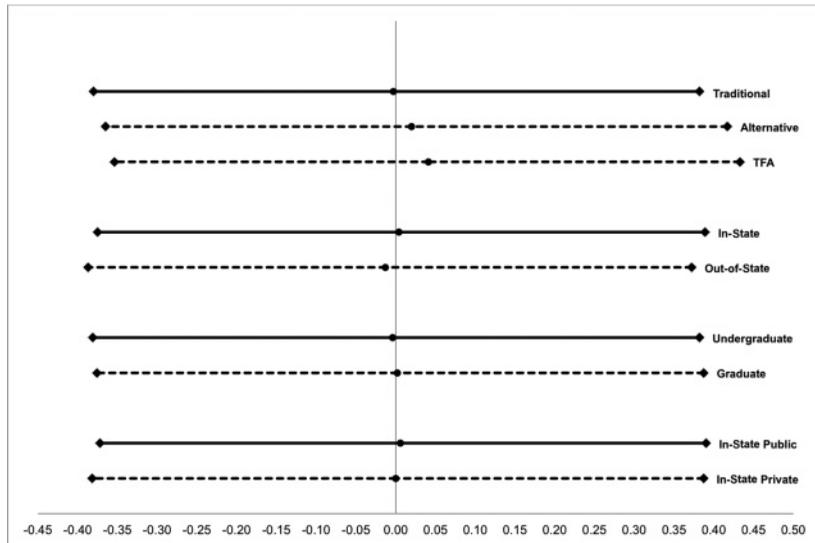
Figure 2. The Distribution of Teacher Effectiveness in Elementary Grades Mathematics. *Note*: Figure depicts the distribution of teacher effectiveness (25th to 75th percentile) for each of our nine teacher categories. Solid lines designate the reference group from value-added models.

standard deviation units) in the student fixed effects models but is no longer statistically significant due to the reduction in case counts (392) and teacher counts (312 traditionally prepared teachers and 19 TFA corps members). Although there were some differences in results across specifications, the overall pattern of findings from the school fixed effect models were supported in our robustness checks.

The Distribution of Teacher Effectiveness
To further contextualize the differences we found between the teacher preparation categories, figures 2 and 3 (elementary mathematics and high school science, respectively) display the mean and interquartile range (twenty-fifth to the seventy-fifth percentile) of the distribution of teacher value-added estimates for each of the teacher preparation categories used in this study.[9] These figures extend and corroborate prior research, illustrating that there is more variation in teacher effectiveness within teacher preparation categories than between them (Boyd et al. 2007; Kane, Rockoff, and Staiger 2008). Despite this general trend there are noteworthy findings regarding the degree of overlap in teacher value-added estimates. For example, figure 3 shows that in relation to their comparison categories, the distributions for TFA, in-state prepared,

---

9. We selected elementary grades mathematics and high school science as examples of the distribution of teacher effectiveness. See Appendix B for teacher effectiveness distribution figures for all five of our teacher preparation category comparisons.
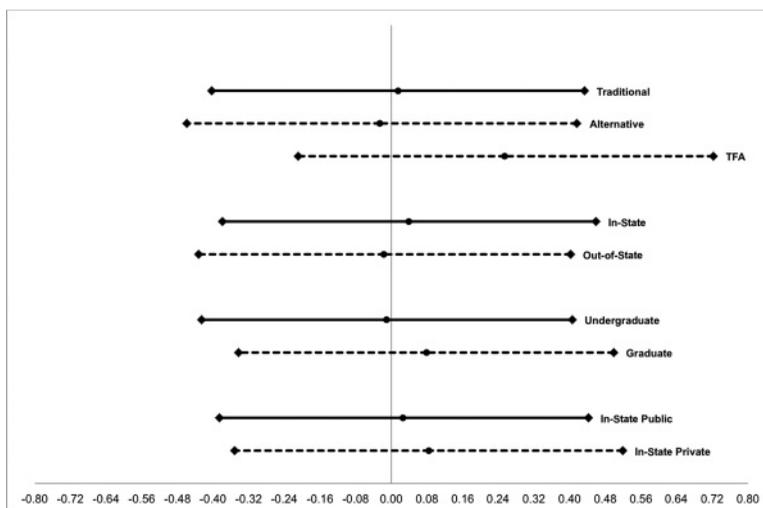
Figure 3. The Distribution of Teacher Effectiveness in High School Science. *Note*: Figure depicts the distribution of teacher effectiveness (25th to 75th percentile) for each of our nine teacher categories. Solid lines designate the reference group from value-added models.

graduate degree prepared, and in-state private university prepared teachers are shifted to the right in high school science. In figure 2 (elementary mathematics), the teacher value-added distributions are more similar, but the significant differences in effectiveness from our analyses are still meaningful due to the number of students being taught by teachers from specific, lower-performing teacher preparation categories (e.g., out-of-state prepared teachers). Overall, results from our first set of analyses indicate that policy choices regarding teacher preparation matter for student achievement.

6. CONCLUSION

All states set policies that affect their educator workforces, but we know little about the extent to which these policies influence student achievement. The most obvious state policies affecting the educator workforce are regulatory—establishing the qualifications needed to teach in public schools. However, a state's financial policies, including incentives for teachers to earn graduate degrees and subsidizing tuition of students enrolled in teacher preparation programs in the state's public institutions of higher education, may also shape the teacher workforce. Overall, we find these regulatory and financial policies, often crafted in reaction to the increased demand for and under-supply of teachers, do affect student achievement in important ways.

Perhaps most striking are the findings that alternative-entry teachers are less effective in high school mathematics and science and out-of-state prepared teachers are less effective in elementary mathematics and reading. These

findings suggest that policies that reduced barriers into teaching may dampen individual student achievement—the practical significance of these effects is amplified by the large numbers of students taught by alternative entry teachers in high school math and science courses and out-of-state teachers in elementary schools. To best craft policy responses to these findings, future research should determine why alternative-entry and out-of-state prepared teachers underperform. For instance, if out-of-state prepared teachers struggle due to difficulties transitioning to the working environment and academic curricula of a new state, induction programs may ease that adjustment. In addition, the findings show little difference in the effectiveness of teachers beginning with undergraduate versus graduate degrees, which suggests that the significant sums of state funds that are used to provide incentives for teachers to obtain graduate degrees may not have the desired effects on student achievement. These findings comparing undergraduate-prepared teachers and graduate-prepared teachers are complex, however, with undergraduate-prepared teachers outperforming graduate-prepared teachers in middle school subjects, but underperforming them in high school science. Future research is needed to explain these contradictory findings. For instance, understanding what types of degrees (education degrees, subject-specific degrees) graduate teachers hold may shed light on our findings. Furthermore, teachers who select into these subjects, and their degrees, may vary in meaningful ways. A more nuanced look at these issues may provide important information about the utility of different traditional teacher preparation pathways and incentives for graduate level education.

It is unlikely that states can or should attempt to turn back the clock and raise barriers into teaching. The demand for more teachers will exceed the supply from the traditional sources without significant innovation in the delivery of instruction or a dramatic change in the incentives to teach in public school classrooms. Instead, it may be helpful to better understand the variations in individuals' skills and abilities and in their preparation to teach that leads to greater effectiveness in the classroom. Here, the findings concerning the effectiveness of TFA corps members may be instructive. TFA selects undergraduates with high cognitive skills as demonstrated by their performance in top colleges and universities *and* based upon noncognitive traits and skills such as motivation, leadership, persistence, and grit. In addition, they provide an orientation to teaching that includes classroom experience during the summer as well as guidance on how to plan for delivery of instruction, manage classrooms, and integrate themselves into schools and communities prior to beginning teaching. TFA also observes corps members and provides feedback multiple times throughout the school year to improve instructional practice. The TFA program itself is relatively small and unlikely to expand to fill a large

amount of the overall demand for teachers, but adding the focus on selection based on noncognitive skills and the methods they use to support new teachers may be beneficial practices to be considered by other teacher preparation programs.

This focus is further underscored by the findings in this study that show there is more variation within the groupings of teachers by entry qualification than between them. This suggests that variations in the skills, abilities, and motivations of those selecting into teaching, in the amount and types of preparation they receive prior to entering teaching, and in the nature and duration of support they receive as they begin their careers, may be able to explain some of the within-group differences. These are the variations we believe are ripe for additional research and could provide additional evidence-based insights about state teacher policies and improving student performance.

REFERENCES

Ashenfelter, Orley. 1978. Estimating the effect of training programs on earnings. *Review of Economics and Statistics* 60(1): 47–57. doi:10.2307/1924332

Baum, Sandy, and Michael S. McPherson. 2011. Sorting to extremes. *Change: The Magazine of Higher Learning* 43(4): 6–12. doi:10.1080/00091383.2011.585289

Bifulco, Robert. 2012. Can non-experimental estimates replicate estimates based on random assignment in evaluations of teacher choice? A within-study comparison. *Journal of Policy Analysis and Management* 31(3): 729–751. doi:10.1002/pam.20637

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management* 24(1): 113–132. doi:10.1002/pam.20072

Boyd, Donald, Pamela Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2006. How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy* 1(2): 176–216. doi:10.1162/edfp.2006.1.2.176

Boyd, Donald, Daniel Goldhaber, Hamilton Lankford, and James Wyckoff. 2007. The effect of certification and preparation on teacher quality. *Future of Children* 17(1): 45–68. doi:10.1353/foc.2007.0000

Boyd, Donald, Pamela Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2009. Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis* 31(4): 416–440. doi:10.3102/0162373709353129

Boyd, Donald, Pamela Grossman, Marsha Ing, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2011. The influence of school administrators on teacher retention decisions. *American Educational Research Journal* 48(2): 303–333. doi:10.3102/0002831210380788

Brunner, Jose Joaquin. 1997. From state to market coordination: The Chilean case. *Higher Education Policy* 10(3–4): 225–237. doi:10.1016/S0952-8733(97)00015-9

Carnoy, Martin, and Susanna Loeb. 2002. Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis* 24(4): 305–331. doi:10.3102/01623737024004305

Clotfelter, Charles, Helen Ladd, and Jacob Vigdor. 2005. Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review* 24(4): 377–392. doi:10.1016/j.econedurev.2004.06.008

Clotfelter, Charles, Helen Ladd, and Jacob Vigdor. 2007. Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review* 26(6): 673–682. doi:10.1016/j.econedurev.2007.10.002

Clotfelter, Charles, Helen Ladd, and Jacob 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. doi:10.1353/jhr.2010.0023

Constantine, Jill, Daniel Player, Tim Silva, Kristin Hallgren, Mary Grider, and John Deke. 2009. An evaluation of teachers trained through different routes to certification, final report. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE 2009–4043).

Council of State Governments (CSG). 2010. *Changing teacher compensation methods: Moving towards performance pay.* Available www.csg.org/policy/documents/TIA_payforperformance_draft2.pdf. Accessed 13 August 2012.

Damon, Amy, and Paul Glewwe. 2011. Valuing the benefits of the education provided by public universities: A case study of Minnesota. *Economics of Education Review* 30(6): 1242–1261. doi:10.1016/j.econedurev.2011.07.015

Darling-Hammond, Linda, Deborah Holtzman, Su Jin Gatlin, and Julian Vasquez Heilig. 2005. Does teacher preparation matter? Evidence about teacher certification, Teach For America, and teacher effectiveness. *Education Policy Analysis Archives* 13(42): 1–48.

Decker, Paul, Daniel Mayer, and Steven Glazerman. 2006. Alternative routes to teaching: The impacts of Teach For America on student achievement and other outcomes. *Journal of Policy Analysis and Management* 25(1): 75–96. doi:10.1002/pam.20157

Dee, Thomas, and Brian Jacob. 2009. The impact of No Child Left Behind on student achievement. NBER Working Paper No. 15531.

Delta Cost Project. 2012. *Trends in college spending (TCS) online.* Available www.tcs-online.org/Reports/Report.aspx. Accessed 10 January 2013.

Dobbie, Will. 2011. Teacher characteristics and student achievement: Evidence from Teach For America. Unpublished paper, Harvard University.

Donaldson, Morgaen, and Susan Moore Johnson. 2010. The price of misassignment: The role of teaching assignments in Teach For America teachers' exit from low-income schools and the teaching profession. *Educational Evaluation and Policy Analysis* 32(2): 299–323. doi:10.3102/0162373710367680

Donaldson, Morgaen, and Susan Moore Johnson. 2011. Teach For America teachers: How long do they teach? Why do they leave? *Phi Delta Kappan* 93(2): 47–51.

Ehrenberg, Ronald G., and Dominic J. Brewer. 1994. Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review* 13(1): 1–17. doi:10.1016/0272-7757(94)90019-1

Feistritzer, Emily C. 2011. *Profile of teachers in the U.S. 2011*. Washington, DC: National Center for Education Information.

Glazerman, Steven, Dan Levy, and David Myers. 2003. Nonexperimental versus experimental estimates of earnings impacts. *Annals of the American Academy of Political and Social Science* 589(1): 63–93. doi:10.1177/0002716203254879

Goldhaber, Dan. 2007. Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources* 47(1): 765–794.

Goldhaber, Dan, and Dominic J. Brewer. 1996. *Evaluating the effect of teacher degree level on educational performance*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Goldhaber, Dan, 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. doi:10.3102/01623737022002129

Goldhaber, Dan, Stephanie Liddle, and Roddy Theobald. 2013. The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review* 34(3): 29–44. doi:10.1016/j.econedurev.2013.01.011

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, Donald, and Tim R. Sass. 2011. Teacher training, teacher quality and student achievement. *Journal of Public Economics* 95(7–8): 798–812. doi:10.1016/j.jpubeco.2010.11.009

Henry, Gary T., and Charles L. Thompson, C. Kevin Fortner, Kevin C. Bastian, and Jade V. Marcus. 2011. Technical report: UNC teacher preparation program effectiveness report. Chapel Hill, NC: Carolina Institute for Public Policy.

Henry, Gary T., Kevin C. Bastian, and Adrienne A. Smith. 2012. Scholarships to recruit the "best and brightest" into teaching: Who is selected, where do they teach, how effective are they and how long do they stay? *Educational Researcher* 41(3): 83–92. doi:10.3102/0013189X12437202

Henry, Gary T., David C. Kershaw, Adrienne A. Smith, and Rebecca A. Zulli. 2012. Incorporating teacher effectiveness into teacher preparation program evaluation. *Journal of Teacher Education* 63(5): 335–355. doi:10.1177/0022487112454437

Ingersoll, Richard, and Lisa Merrill. 2010. Who's teaching our children? *Educational Leadership* 67(8): 14–21.

Johnstone, D. Bruce. 2003. Cost sharing in higher education: Tuition, financial assistance, and accessibility in a comparative perspective. *Czech Sociological Review* 39(3): 351–374.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review* 27(6): 615–631. doi:10.1016/j.econedurev.2007.05.005

Ladd, Helen. 2011. Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement? *Educational Evaluation and Policy Analysis* 33(2): 235–261. doi:10.3102/0162373711398128

Lankford, Hamilton, Susannah 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. doi:10.3102/01623737024001037

Long, Bridget T. 2004. Does the format of a financial aid program matter? The effect of state in-kind tuition subsidies. *Review of Economics and Statistics* 86(3): 767–782. doi:10.1162/0034653041811653

Luebke, Bob. 2011. *We need a better way to pay teachers?* Available www.nccivitas.org/2011/we-need-a-better-way-to-pay-teachers/. Accessed 30 January 2012.

McCaffrey, Daniel F., J. R. Lockwood, Daniel Koretz, Thomas A. Louis, and Laura Hamilton. 2004. Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics* 29(1): 67–101. doi:10.3102/10769986029001067

National Research Council (NRC). 2010. *Preparing teachers: Building evidence for sound policy.* Washington, DC: National Academies Press.

North Carolina General Assembly (NCGA). 1985. *North Carolina general statutes: Chapter 115C: Elementary and secondary education.* Available www.ncga.state.nc.us/EnactedLegislation/Statutes/HTML/ByArticle/Chapter_115C/Article_20.html. Accessed 8 April 2013.

North Carolina General Assembly (NCGA). 2012. *Excellent Public Schools Act: Session 2011.* Available www.ncga.state.nc.us/sessions/2011/bills/senate/PDF/S795v0.pdf. Accessed 9 April 2013.

Poterba, James M. 1996. Government intervention in the markets for education and health care: How and why? In *Individual and social responsibility: Child care, education, medical care, and long-term care in America,* edited by Victor R. Fuchs, pp. 277–308. Chicago: University of Chicago Press.

Raymond, Margaret, Stephen Fletcher, and Javier Luque. 2001. *Teach For America: An evaluation of teacher differences and student outcomes in Houston, Texas.* Stanford, CA: CREDO, Stanford University.

Reininger, Michelle. 2012. Hometown disadvantage? It depends on where you're from: Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis* 34(2): 127–145. doi:10.3102/0162373711420864

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. Teachers, schools, and academic achievement. *Economometrica* 73(2): 417–458. doi:10.1111/j.1468-0262.2005.00584.x

Ronfeldt, Matthew, Susanna Loeb, and James Wyckoff. 2013. How teacher turnover harms student achievement. *American Educational Research Journal* (50)1: 4–36.

Rothstein, Jesse. 2010. Teacher quality in education production: Tracking, decay, and student achievement. *Quarterly Journal of Economics* 125(1): 175–214. doi:10.1162/qjec.2010.125.1.175

Roza, Marguerita, and Raegen Miller. 2009. *Separation of degrees: State-by-state analysis of teacher compensation for master's degrees.* Seattle, WA: Center for American Progress.

Sass, Tim R. 2011. Certification requirements and teacher quality: A comparison of alternative routes to teaching. CALDER Working Paper No. 64.

Schneider, Mark, and Jorge Klor de Alva. 2011. *Cheap for whom? How much higher education costs taxpayers. Education Outlook No. 8.* Washington, DC: American Enterprise Institute for Public Policy Research.

Shadish, William, M. H. Clark, and Peter Steiner. 2008. Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association* 103(484): 1334–1343. doi:10.1198/016214508000000733

Winston, Gordon C. 1999. Subsidies, hierarchy, and peers: The awkward economics of higher education. *Journal of Economic Perspectives* 13(1): 13–36. doi:10.1257/jep.13.1.13

Wooldridge, Jeffrey M. 2009. *Introductory econometrics: A modern approach.* Mason, OH: South-Western Cengage Learning.

Xu, Zeyu, Jane Hannaway, and Colin Taylor. 2011. Making a difference? The effects of Teach For America in high school. *Journal of Policy Analysis and Management* 30(3): 447–469. doi:10.1002/pam.20585

APPENDIX A: SPECIFICATION CHECKS

Table A.1. Comparisons of Average Effectiveness for Policy Relevant Preparation Categories (Rich Covariate OLS Model)

Policy Relevant Comparison	Elementary Math	Elementary Reading	Middle Math	Middle Reading	High School Math	High School Science	High School English	High School Social Studies
Alternative vs. Traditional	0.015	−0.001	−0.015*	−0.003	−0.052**	−0.043**	−0.006	−0.052**
	(0.010)	(0.007)	(0.007)	(0.004)	(0.011)	(0.015)	(0.007)	(0.014)
Student Observations Used	431,784	607,106	253,864	282,680	233,062	158,411	127,903	188,843
TfA vs. Traditional	0.055**	0.034*	0.127**	0.029**	0.213**	0.296**	0.042*	0.129**
	(0.020)	(0.016)	(0.020)	(0.011)	(0.031)	(0.045)	(0.018)	(0.042)
Student Observations Used	431,784	607,106	253,864	282,680	233,062	158,411	127,903	188,843
Out-of-State vs. In-State	−0.024**	−0.011**	0.002	−0.003	−0.024	−0.070**	−0.005	−0.023
	(0.005)	(0.004)	(0.008)	(0.005)	(0.014)	(0.020)	(0.009)	(0.016)
Student Observations Used	403,502	564,390	164,850	178,069	161,778	69,927	85,901	143,373
Graduate Degree vs. Undergraduate	0.005	0.006	−0.007	−0.014*	−0.003	0.068**	0.006	0.037
	(0.008)	(0.006)	(0.014)	(0.006)	(0.019)	(0.022)	(0.009)	(0.019)
Student Observations Used	403,502	564,390	164,850	178,069	161,778	69,927	85,901	143,373
In-State Private vs. In-State Public	−0.006	−0.004	−0.012	0.007	−0.004	0.069	0.003	−0.009
	(0.007)	(0.005)	(0.012)	(0.008)	(0.018)	(0.036)	(0.010)	(0.017)
Student Observations Used	238,981	342,521	98,388	100,877	117,032	38,851	64,557	99,018

Notes: In these analyses the second category in each row is the reference category (traditionally-prepared, in-state-prepared, undergraduate-degree prepared, and in-state public-university prepared teachers). *Student Observations Used* indicates the number of cases used in these OLS models.
*Statistically significant at the 5% level; **statistically significant at the 1% level.

Table A.2. Comparisons of Average Effectiveness for Policy Relevant Preparation Categories (Student Fixed Effects Model)

Policy Relevant Comparison	Elementary Math	Elementary Reading	Middle Math	Middle Reading	High School Math	High School Science	High School Social Studies
Alternative vs. Traditional	−0.017	0.010	−0.020	0.009	0.004	−0.084	−0.015
	(0.016)	(0.012)	(0.018)	(0.011)	(0.024)	(0.047)	(0.032)
Student Observations Used	10,668	17,202	19,943	25,394	20,227	6,507	10,467
TFA vs. Traditional	0.036	0.017	0.143**	0.020	0.128*	0.142	0.041
	(0.030)	(0.022)	(0.049)	(0.028)	(0.061)	(0.101)	(0.099)
Student Observations Used	2,136	3,622	1,254	2,284	1,807	392	709
Out-of-State vs. In-State	−0.024**	−0.008	−0.008	0.003	−0.051	−0.152	−0.040
	(0.008)	(0.005)	(0.024)	(0.018)	(0.030)	(0.086)	(0.051)
Student Observations Used	42,333	57,957	10,107	11,137	12,725	1,453	6,458
Graduate Degree vs. Undergraduate	0.008	0.012	−0.069	−0.020	−0.017	0.067	0.003
	(0.012)	(0.009)	(0.042)	(0.023)	(0.037)	(0.085)	(0.056)
Student Observations Used	17,597	23,428	3,225	6,337	7,920	1,215	5,376
In-State Private vs. In-State Public	−0.008	−0.012	−0.018	0.004	0.027	0.079	−0.059
	(0.014)	(0.010)	(0.046)	(0.037)	(0.043)	(0.185)	(0.060)
Student Observations Used	13,560	21,140	2,240	2,057	5,007	333	2,467

Notes: In these analyses the second category in each row is the reference category (traditionally prepared, in-state prepared, undergraduate degree prepared, and in-state public university prepared teachers). *Student Observations Used* indicates the number of students experiencing within-unit (student) variation for the comparison.
*Statistically significant at the 5% level; **statistically significant at the 1% level.

Table A.3. Unique Teacher Counts from Value-Added Models

Teacher Preparation Categories	Elementary Math	Elementary Read	Middle Math	Middle Read	High School Math	High School Science	High School English	High School Social Studies
Traditional								
OLS	7,227	7,329	1,584	1,841	1,152	577	817	981
School FE	2,931	2,989	1,209	1,434	946	421	506	535
Student FE	1,280	1,355	788	954	710	312	—	418
Alternative								
OLS	577	592	943	1,176	720	741	505	372
School FE	566	580	812	1,001	629	535	386	309
Student FE	416	435	560	697	493	331	—	249
TFA								
OLS	115	119	86	118	71	56	49	38
School FE	109	112	74	96	52	32	31	33
Student FE	98	102	51	77	43	19	—	26
In-State								
OLS	4,062	4,130	916	1,011	790	329	578	655
School FE	3,148	3,191	550	663	471	140	245	360
Student FE	2,201	2,259	389	448	375	89	—	265
Out-of-State								
OLS	3,165	3,199	608	830	362	248	239	326
School FE	2,992	3,029	563	701	299	136	164	240
Student FE	2,212	2,245	373	450	245	88	—	200
Undergraduate								
OLS	6,491	6,581	1,441	1,585	982	421	647	776
School FE	3,043	3,102	391	612	382	132	216	308
Student FE	1,608	1,630	215	351	292	83	—	234
Graduate								
OLS	736	748	143	256	170	156	170	205
School FE	726	739	124	299	151	106	123	173
Student FE	602	621	101	159	133	72	—	143
In-State Public								
OLS	2,858	2,910	739	822	633	264	465	510
School FE	1,603	1,630	197	220	237	53	136	163
Student FE	910	940	112	113	204	26	—	131
In-State Private								
OLS	1,204	1,220	177	189	157	65	113	145
School FE	1,053	1,072	125	124	130	33	82	106
Student FE	714	732	84	68	115	19	—	81

Note: This table displays unique teacher counts from OLS, school fixed effect, and student fixed effect value-added models. Because there is only one high school English exam (English I) teacher counts are unavailable for student fixed effects.
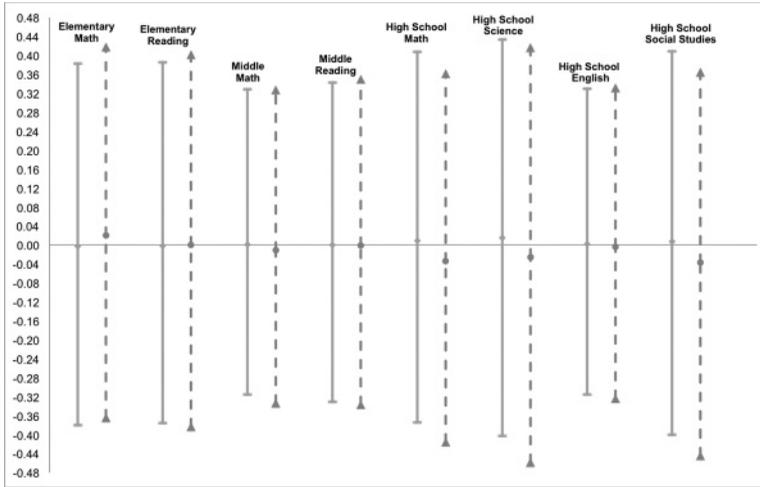FE: fixed effects.

APPENDIX B

Figure B.1. The Distribution of Teacher Effectiveness for Traditionally Prepared vs. Alternative Entry Teachers. *Note*: Lines for traditionally prepared teachers are continuous and on the left. Lines for alternative entry teachers are dashed and on the right.
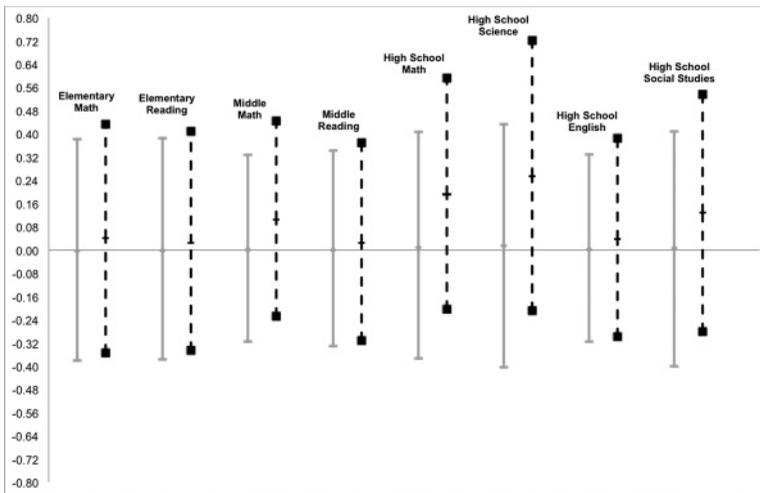

Figure B.2. The Distribution of Teacher Effectiveness for Traditionally Prepared vs. Teach For America Corps Members. *Note*: Lines for traditionally prepared teachers are continuous and on the left. Lines for TFA corps members are dashed and on the right.
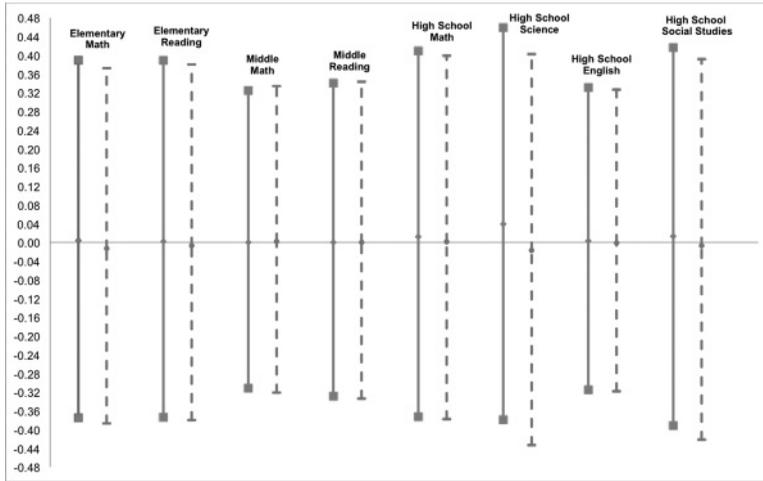
Figure B.3. The Distribution of Teacher Effectiveness for In-State Prepared vs. Out-of-State Teachers. *Note*: Lines for in-state prepared teachers are continuous and on the left. Lines for out-of-state prepared teachers are dashed and on the right.
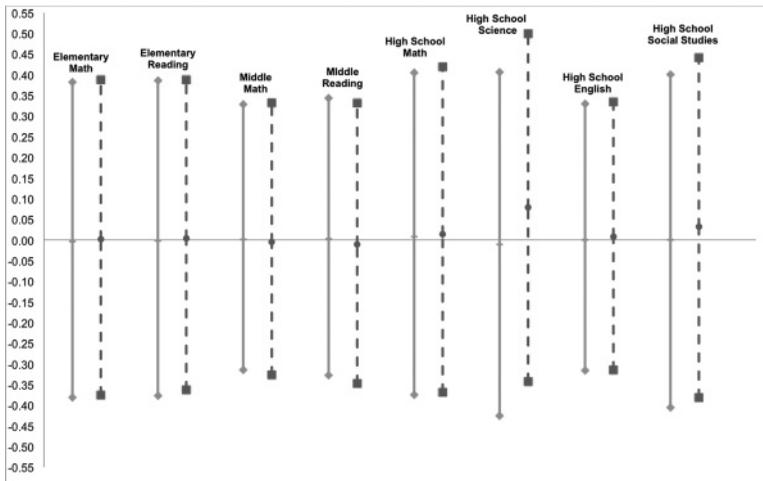

Figure B.4. The Distribution of Teacher Effectiveness for Undergraduate-Degree Prepared vs. Graduate-Degree Prepared Teachers. *Note*: Lines for undergraduate-degree prepared teachers are continuous and on the left. Lines for graduate-degree prepared teachers are dashed and on the right.
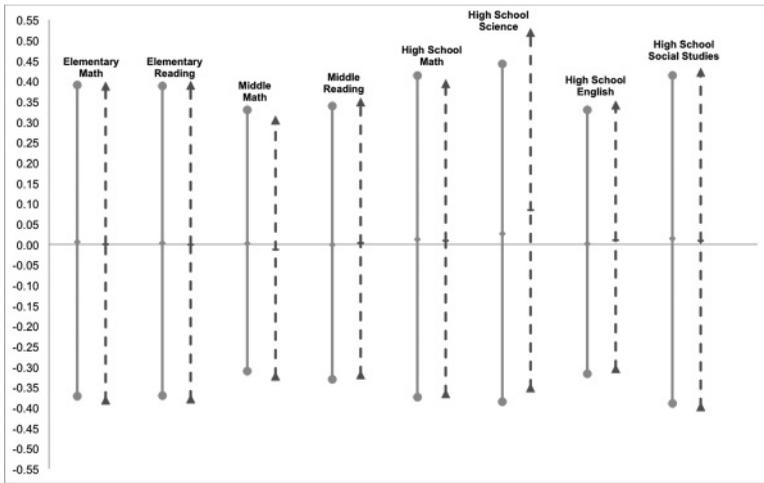
Figure B.5. The Distribution of Teacher Effectiveness for In-State Public-University Prepared vs. In-State Private-University Prepared Teachers. *Note*: Lines for in-state public-university prepared teachers are continuous and on the left. Lines for in-state private-university prepared teachers are dashed and on the right.