Teach for America and Teacher Quality: Increasing Achievement over Time

Emily K. Penner*
Assistant Professor
School of Education
University of California, Irvine

* Please direct correspondence to:
emily.penner@uci.edu
3200 Education
University of California, Irvine
Irvine, CA 92697

Emily K. Penner is an Assistant Professor of Education at the University of California, Irvine. Her research focuses on educational inequality and policy, and considers the ways that districts, schools, teachers, and parents can contribute to or ameliorate educational inequality.

Acknowledgements: I would like to thank Kara Bonneau and the North Carolina Education Research Data Center and their invaluable assistance with providing administrative data from North Carolina. I would also like to acknowledge the support of Jason Atwood, Alissa Swartz and the External Research Partnerships team at Teach For America and their assistance in identifying Teach For America corps members in North Carolina. I am grateful to Greg J. Duncan, Marianne Bitler, Deborah Lowe Vandell, George Farkas, Thurston Domina, and participants at the Annual Meetings of the Sociology of Education Association and the Society for Research on Educational Effectiveness for their valuable comments.

Research reported in this study is supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P01HD065704. This research is also supported by the Society for Research in Child Development Dissertation Funding Award. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health or the Society for Research in Child Development.

Teach for America and Teacher Quality: Increasing Achievement over Time

Abstract: Teach For America (TFA), a prominent alternative teacher certification program, has evolved substantially since it began in 1989. However, most studies ignore this variation when evaluating TFA's effects on student outcomes. Using twelve years of administrative data, I test whether TFA's effects on achievement differ over time, focusing on a major reform introduced in 2005 and testing for changes in the qualifications of TFA corps members recruited before and after this reform. I find that TFA's effects on achievement increased over time in some subjects and grade levels. However, these changes do not occur immediately following nor are they explained by most observable teacher quality indicators. Instead, results indicate gradual improvement over time in many subjects. These results help explain divergent findings in previous research from different periods, and underscore the importance of understanding the continuing evolution of alternative teacher certification programs.

Teach For America (TFA) is a selective alternative teacher certification program that has influenced conversation and policy about how teachers are recruited and trained in schools across the country. TFA teachers represent a small fraction of teachers in American primary and secondary classrooms nationwide, even in their largest placement districts where they make up at most three percent of teachers.[i] Yet TFA's prominence in the national dialogue around teacher recruitment and certification, and its continued expansion to low-income districts across the country suggests that TFA's effects should be carefully examined.

As TFA's primary aim for many years was to close achievement gaps, a growing body of work has evaluated TFA's effects on student achievement at different points in time. Early evidence on TFA's effects was inconclusive, and of mixed quality, but a number of recent studies find positive effects of TFA in some subjects and grades. Many of these evaluations treat TFA as a singular program that was just as effective in its early years as it is more recently, but TFA sees itself as a dynamic organization that has evolved over time to improve the effectiveness of its teachers. Among many changes, TFA refined its training curriculum based on five leadership traits in the early 2000s, it completed an initial push for substantial expansion in 2005, and it began incorporating analysis of student data to evaluate corps member performance in 2005 (Jaramillo, 2002; Sawchuk, 2009; Tourangeau, 2003). Subsequently, TFA initiated a second period of expansion into the mid-2010's (Kopp, 2011). However, research focused on its effects on achievement has largely ignored whether TFA teacher impacts changed over time as TFA modified its program in an effort to improve student outcomes. Drawing on work documenting that TFA is not uniformly effective in all subjects and grade levels, this study examines whether TFA's impact on student achievement has changed over time, paying specific attention to how changes may vary across subjects and grade levels.

Specifically, this study examines the effect of TFA on achievement across elementary, middle, and high school using a unique dataset from the state of North Carolina, spanning the years 1999/2000 – 2010/2011. These data allow me to test whether TFA's impacts on achievement have changed across the pre- and post-2005 period, as North Carolina experienced an expansion that mirrored TFA's national expansion. In this study I thus: 1) explore the differential effects of TFA teachers on achievement across subjects and grade levels, 2) examine whether these effects change over time, and 3) investigate the degree to which differences in teacher qualifications can account for any differences in TFA's effects on achievement or changes in these effects over time.

I find evidence that TFA's effects on achievement are sustained over time and increased in several subjects and grade levels, specifically in middle school math, high school science, and high school social studies. This pattern of improvement is shared more broadly across most subjects and grades, although the improvements are only statistically significant for a subset of areas. In some of those cases, TFA had a positive impact early on and its impact grew over time. In others, TFA's impact was initially negative and improved substantially in later years. However, most of these changes do not occur immediately after program reforms in the mid-2000s, but rather develop over time. Nor are the changes a result of most observable indicators of teacher qualifications. This suggests that TFA's sustained effects on student achievement and improvements over time are likely due to a combination of factors that contribute to a trend of incremental, positive improvements, some of which are not measurable in existing administrative data.

Literature review

Teach For America: Background

From its inception, TFA challenged the notion that effective teachers need extensive training or traditional certification before entering the classroom (Kopp, 2011). Instead, TFA asserts that individuals from backgrounds with high degrees of academic rigor, leadership experience, determination, and dedication to the mission of providing children with an excellent education can improve student outcomes with minimal training or teaching experience. TFA also promotes the idea that teachers who can succeed in improving student outcomes can be identified through a rigorous selection process that evaluates personal qualities without evidence of specific teaching skill.

Prior experimental and quasi-experimental evidence suggests that TFA teachers have a positive impact on student achievement compared with other math and science teachers, but not reading or English teachers at the elementary and high school levels (Chiang, Clark, & McConnell, 2017; Clark et al., 2013; Glazerman, Mayer, & Decker, 2006; Kane, Rockoff, & Staiger, 2008; Xu, Hannaway, & Taylor, 2011). Experimental evidence from Clark et al. (2015) suggests that TFA teachers are as effective as non-TFA teachers in raising student achievement in both reading and math, with the exception of lower-elementary reading grades where TFA teachers were slightly better at raising achievement than non-TFA teachers. Other quasi-experimental and descriptive studies using administrative data yield mixed conclusions, with some finding positive achievement effects of TFA teachers relative to non-TFA teachers, others finding a varied pattern by subject, and still others finding that TFA teachers perform no differently or worse than non-TFA teachers (Darling-Hammond, Brewer, Gatlin, & Vasquez Heilig, 2005; Schoeneberger, Dever, & Tingle, 2009; Ware et al., 2011). However, some of the studies that find negative effects of TFA compare TFA to teachers across a broad set of districts and regions. The non-TFA teachers in these studies are often teaching populations of relatively

advantaged students, rather than looking exclusively at teachers in the same under-resourced schools where TFA places teachers (c.f. Darling-Hammond et al., 2005; Laczko-Kerr & Berliner, 2002).

The literature examining TFA's effects on non-test outcomes is still emerging. Several studies have examined the effects of TFA on student absences and suspensions, yielding conflicting results. Backes and Hansen (2018) find suggestive evidence that students of TFA teachers had fewer absences and suspensions relative to students in the same schools with non-TFA teachers, while Decker et al. (2004) and Clark et al. (2013) find no evidence of any differences. Backes and Hansen (2018) also examined the effects of TFA on several other non-test outcomes, including GPA, course failure, and grade repetition, finding no evidence of TFA effects. Author (2014) examined the effects of TFA on high school completion and plans for higher education, finding positive effects on completion, but negative effects on higher education plans, which is explained by their successful efforts to help more lower-performing students complete high school.

A growing body of research has also critiqued the TFA model and its impact on schools and districts serving marginalized students. Some find evidence of conflicts between TFA's training model and equity-oriented teaching and school management (McNew-Birren, Hildebrand, & Belknap, 2018; Trujillo, Scott, & Rivera, 2017). Others argue that TFA displacement within schools and subjects contributes to higher turnover among TFA corps members than other new teachers (Donaldson & Johnson, 2010). The harshest critiques of TFA have come from those who examine TFA's often close relationship with school choice and charter schools and their heavy reliance on "no-excuses" discipline techniques to buttress what critics argue is an excessive focus on academic performance, narrow standards, and less-

regulated, corporate-run education (Brewer and Wallis 2015; Brewer et al., 2016; Kretchmar, Sondel, & Ferrare, 2018; Kumashiro 2010; Lahann and Reagan 2011; Ledebre & Thomas 2017; Mungal 2016; Sondel, 2015; Trujillo, Scott, & Rivera, 2017).

Researchers who find both supportive and critical evidence about TFA typically treat it as a static intervention. Much of the evidence of TFA impacts examines TFA's effects on achievement and other outcomes during only a short time period. However, TFA as an organization has undergone many substantial transitions since its inaugural year in 1989-1990. TFA underwent a period of expansion in the early 2000s, expanding from five original sites, including one in Eastern North Carolina, to 48 regional sites (with approximately 11,000 corps members) in the 2013-14 school year (teachforamerica.org, 2013c).

In addition to expanding its national presence, TFA has also worked to respond to critics of the program by reorganizing and improving its teacher-training program and recruitment. After several high-profile challenges (e.g., Darling-Hammond, 1994), TFA strove to strengthen the quality of its training program in the late 1990s. Its first response was to focus more heavily on five core leadership traits in its national training and mentoring programs. It also sought to build its capacity to evaluate its own effectiveness by tracking student progress internally, which was also incorporated into its teacher-training model in 2005 (teachforamerica.org, 2016). In this training reform, TFA redoubled its focus on improving student academic performance. Early on in its existence, TFA focused on "closing the achievement gap" for students in the schools it serves, and put a large stake in promoting, "significant gains," (meaning 1.5 to 2 grade levels of improvement in core subject areas) for its students. For many of the students it serves, this target aimed to catch them up to grade level standards. It has since expanded its mandate to "alter educational trajectories" (teachforamerica.org, 2013a) which it hopes to achieve by promoting a

passion for learning and school engagement, but TFA remains largely focused on improving student achievement (Farr, 2010). Finally, in more recent years, the organization has worked to address critiques of TFA's culture by increasing the racial and socio-economic diversity of its corps, by recruiting more individuals with graduate degrees and post-baccalaureate professional experience, and by considering student growth and opportunities beyond achievement (Jennings, 2014; teachforamerica.org, 2013c).

However, few studies have examined changes in TFA's effects over time and only for elementary and middle school grades in New York City (Boyd, Dunlop, et al., 2012). Boyd and colleagues look over ten years, and find little indication of consistent improvements in TFA's effects on student achievement over time relative to veteran teachers trained through college-recommended teacher preparation programs. However, there is reason to suspect that the New York teaching context might differ from other TFA sites as New York has experienced a substantial narrowing of the gap in teacher quality between poor and non-poor schools due to overall quality improvements, while evidence elsewhere suggests the persistence of such gaps in quality (Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008; Clotfelter, Ladd, Vigdor, & Wheeler, 2006; Henry et al., 2014). This work also excludes high school, where the largest TFA effects have been found in other prior studies (Henry et al., 2010, 2014).

Teacher qualifications and student achievement across subjects and grade levels

The tendency to describe TFA as a static intervention extends to the conceptualization of TFA's corps members as individuals with homogeneous backgrounds and similar levels of preparation through the evolution of the program. TFA's more recent efforts to diversify its corps members and recruit from a wider set of colleges and universities, though questioned, suggest that this is not the case (Irizarry & Donaldson, 2012; White, 2016). Given the varied

relationships between measures of teacher qualifications and student achievement, and the myriad differences between TFA and other teachers, we might expect teacher qualifications to help explain any differences in the effects TFA teachers have on achievement overall, but also any changes that might have occurred as the program has evolved.

The prior literature examining the relationship between experience and qualifications suggests that some qualifications matter consistently while others do not. For example, the literature examining the relationship between teacher experience and student achievement indicates a positive relationship (Clotfelter, Ladd, & Vigdor, 2007, 2010; Ladd & Sorensen, 2017; Papay & Kraft 2015), while many other observable measures of teacher qualifications, such as having a master's degree, are inconsistently related to teacher's achievement effects (Chingos & Peterson 2011; Clotfelter et al., 2010).

An alternative source of temporal variation in the effects that TFA teachers have on student achievement are changes in the relative qualifications of non-TFA teachers who are willing to work in TFA placement schools. Any potential changes in the backgrounds of these teachers might explain changes in TFA's effects and might lead principals to try to recruit more or fewer TFA teachers over time. Research about the general pool of teachers suggests that teachers are coming from weaker academic backgrounds over several decades. Prospective education majors have below-median average SAT scores (Ballou & Podgursky, 1997), and are over-represented in the bottom two SAT quintiles (Vance & Schlechty, 1982). Although average teacher qualification, as measured by the SAT and ACT end of high school aptitude tests, has declined only slightly between the 1950s and the 2000s, the percentage of top-of-class females entering the teaching force declined substantially over the same time period (Corcoran, Evans, & Schwab, 2004). Evidence from North Carolina in more recent years confirms that teachers in

high-poverty schools, including those where TFA teachers are placed, are more inexperienced, attended less competitive undergraduate institutions, with below-mean licensure test scores, and are less likely to be Nationally Board Certified than teachers elsewhere in the state (Clotfelter et al., 2007; Clotfelter et al., 2006).

In sum, the previous literature on the impacts of observable measures of teacher qualification suggests that factors such as years of experience and certification test scores are likely to have a larger impact on student achievement than other measures of teacher qualification, both for TFA and non-TFA teachers. Likewise, these observable measures of teacher qualifications seem the most likely to be able to explain any of the effects of TFA or any changes in its effects over time. It also seems likely that controlling for measures of teacher qualification in which TFA is relatively weak, such as experience and having a master's degree, might increase TFA's effects relative to counterfactual teachers, but that this impact might decline over time as more TFA alumni stay in teaching and gain experience and additional credentials.

Data

I examine the relationship between TFA, student achievement, and observable teacher qualification over time using administrative data from the state of North Carolina. TFA places teachers in three regions in North Carolina. Eastern North Carolina, which encompasses rural and urban areas, opened in 1990 and was one of TFA's charter locations; Charlotte, which is a primarily urban region, has had corps members since 2004; and North Carolina Piedmont Triad, which placed its first corps members in 2015 and is not part of this study (Teach For America, 2017).

The North Carolina Education Research Data Center (NCERDC) at Duke University

provided administrative data from North Carolina. NCERDC provided demographic and assessment records for students matched to their classroom teachers.[ii] These data are well suited to the examination of the effects of TFA versus non-TFA teachers on achievement. In addition to prior studies that directly compare TFA teachers to non-TFA teachers (Henry et al., 2014; Xu et al., 2011), these data have also been used to examine teacher quality (Clotfelter et al., 2006).

All North Carolina TFA teachers (called corps members) from the entering corps of 1999 to 2010[iii] were identified with the assistance of the External Research Partnerships Team at TFA. 1,502 individuals were identified by TFA as being assigned to the two North Carolina regions. These individuals were then matched to NCERDC files based on a variety of characteristics including social security number, first and last name, assigned school name, and assigned Local Education Agency (LEA).[iv] Of the 1,502 unique TFA observations, 1,304 were successfully matched to a teacher observation by NCERDC, and 904 of these teachers were assigned to teach a tested subject.[v] Of the successfully matched TFA corps members, 699 were placed in Eastern North Carolina between 1999/2000 and 2010/2011 and 605 were placed in Charlotte between 2004/2005 and 2010/2011. In addition to the demographic characteristics used to match TFA teachers to teacher records in the NCERDC database, this information also included regional placement and whether the individuals completed their two year commitment or not. The majority of these individuals were placed in North Carolina after 2004, when the Charlotte region was added and number of new corps members placed in Eastern North Carolina more than doubled. TFA corps members by placement region and year for grades K-12 in 1999/2000 – 2010/2011 are shown in Table 1. Given that only the Eastern North Carolina region had corps members through much of the early period, I also examine patterns separately for this region. As I find similar results, I do not report them separately here.

[Insert Table 1 Here]

From 1999/2000-2010/2011, TFA teachers taught in 435 schools in 74 LEAs and charter schools in North Carolina, primarily concentrated in the northeastern corner of the state or near the Charlotte metro area. I concentrate on the TFA and non-TFA teachers in tested subjects in the same schools, grades, and years (or schools, subjects, and years in high school). This allows me to compare the scores of students with TFA teachers to scores of students who might have been assigned to a TFA teacher but had a non-TFA teacher for the same course (or grade level). I describe this approach in more detail in the methods section below, but to provide evidence about how students and teachers in these two groups differ, a comparison of selected demographic characteristics of the students with a TFA teacher versus four different comparison groups is shown in Table 2. The four comparison categories are: (1) students with non-TFA teachers in the same schools, grades, and years as students in TFA classrooms; (2) students in the same schools as TFA teachers, but in different grades and/or years; (3) students in LEAs that have TFA teachers, but in schools that never had a TFA teacher; and (4) students in LEAs that never had a TFA teacher. The first of these is the best comparison, because these students had the potential to be assigned to a TFA teacher in the same year. I focus on this comparison in the analyses described below. The analytic sample throughout this study consists of only those students with TFA teachers (shown in column 1), and those who had the potential of being assigned to a TFA teacher in comparison category 1 (shown in column 2).

[Insert Table 2 here]

As indicated in Table 2, students of TFA teachers are most similar to other students in non-TFA classrooms in the same schools across a wide variety of characteristics. Table 2 also reveals large differences in the observable teacher qualification measures across the five groups.

TFA teachers are much less experienced (less than one year on average compared with more than 11 in all of the other groups), less likely to be fully certified (less than half compared with three-fourths of non-TFA teachers), and less likely to have a master's degree (fewer than five percent compared with at least 25 percent). However, TFA teachers' mean certification scores (Praxis scores) are .4 standard deviations higher than those of the counterfactual teachers in the same schools, grades (subjects), and years in which TFA teachers are teaching.[vi] This suggests that while TFA teachers lack the experience and formal credentials of non-TFA teachers, they may have a stronger grasp of the material the certification board deemed important for teaching, despite limited formal training.

The outcomes of interest in this study are end-of-grade or end-of-course student test scores. Students in North Carolina are tested beginning in grade three, but in order to include a lagged test score control, I examine outcomes beginning in grade four. Tested subjects in elementary and middle school include math and reading, which I combine across grades four to five and six to eight for parsimony.[vii] Students are tested in four subjects in high school: math, English, science, and social studies. Three of these four consist of different subject area that I combine in the primary models shown below. High school math consists of Algebra 1, Algebra II, and Geometry; high school science consists of Chemistry, Biology, Physical Science, and Physics; social studies consists of US History, Civics and Economics, and Economy, Legal, and Political Systems; and English consists of English 1.[viii] Not all of the individual subjects are tested each year, but at least one subject in each global subject area was tested every year, with the exception of 2005, when no social studies was test given. Student test scores are standardized within grade and year for end-of-grade tests and within year and subject for high school end-of-course tests.[ix] Because the test score means and standard deviations presented in Table 2

aggregate across grades three through eight and nine through twelve respectively, the values in each row do not show an exact mean of zero and standard deviation of one. Lagged test score controls are included in models examining outcomes in grades four through eight.[x] As students do not take a consistent course sequence in the same grades across the state, or even within schools, models examining high school outcomes include controls for eighth grade language arts and math scores. All models include additional terms for squared and cubed prior achievement scores.

Models also include several student control variables. Student gender is coded as a 1 if the student is female and a 0 if the student is male. NCERDC identifies six mutually exclusive race/ethnicity categories: white, black, Hispanic, Asian, American Indian, and other ethnicity. For parent education, the maximum value across all of the years the student was in the data was coded into three categories: a high school diploma or less, some college or trade school, and a college degree or more. For English proficiency status, disability status, gifted status, and ever repeated a grade, a student was given a value of 1 for these variables if this was ever indicated as true across all of the years the student was present in the data, and a 0 if not. I also calculated indicators for being over- and under-age for grade.[xi] The student over-and under-age variables change following a retention or early promotion, unlike the rest of the student characteristics that are filled in across multiple years of student observations.

North Carolina began compiling a database of teacher qualification measures in 1995. Thus, although the student data begin in 2000, I include teacher data beginning in 1995. The majority of the observable teacher qualification measures are allowed to vary across years, with the exception of variables that are collected as a teacher is first hired in North Carolina. Years of experience is a continuous variable, which changes over time, that is taken from a teacher's pay

record and documents the number of years a teacher has occupied their current position. From this variable, I also create an indicator for if the teacher is a "veteran" teacher with four or more years of experience, following Glazerman et al. (2006). I also include indicators for whether a teacher is fully certified and whether they have a master's degree or higher (yes = 1; no = 0). I create a variable for undergraduate selectivity by matching students' undergraduate degree granting institutions to the 2009 Barron's Selectivity rankings, which include four categories (most competitive; highly competitive plus; highly competitive; very competitive). The remainder of the universities are included in a fifth category of unranked universities, which serves as the reference category in the analyses that examine teacher qualifications. The final teacher qualification measure is average Praxis licensure test score. Praxis scores are standardized within test type and the year in which the test was taken. These standardized values are then averaged within the individual teacher.

By filling in modal and maximal values using data from multiple datasets provided by the NCERDC, I am able to reduce the amount of missing data substantially. However, some variables still have a substantial amount of missing data, particularly for a few of the observable teacher qualification measures and lagged student test score controls. To include the maximum number of observations possible in the analyses, I recode the missing values to zero and include an indicator for if the variable is missing.

Method

To examine the relationship between having a TFA teacher and student achievement, I estimate models with fixed effects for school, grade, and year combinations.[xii] The school-grade-year fixed effects approach allows me to compare students in TFA and non-TFA classrooms within the same school, grade, and year. This approach follows Boyd et al. (2006), who examine

how achievement gains differ by teacher training pathways in New York City in grades 4-8. This type of comparison hinges on the possibility that a given student could have a TFA teacher. Because TFA only places corps members in two regions in NC, students in many NC schools and LEAs have no chance of having TFA teacher. Thus, the evaluation of achievement gains in TFA versus non-TFA classrooms is restricted to schools, grades, and years with at least one TFA and one non-TFA teacher. This is important because TFA teachers are not randomly assigned to schools throughout the state or even within a given LEA, resulting in the substantial differences between non-TFA students in schools and LEAs with and without TFA teachers (see Table 2).

To examine variation in the association between having a TFA teacher and student achievement across grade levels and subject areas, the analyses use variations on the general dynamic panel data model, adapted from Boyd et al. (2006). The reduced form model is as follows:

$$A_{igsy} = \gamma_o + \gamma_1 A_{igs,y-1} + \gamma_2 S_i + \gamma_3 TFA_{iy} + \pi_{sgy} + \varepsilon_{isgy} \quad (1)$$

In this model, the standardized achievement level, A_{igsy} from the End-of-Grade or End-of-Course standardized test score of student i, in grade g, in school s, in year y, is a function of the student's prior-year test score (including linear, squared, and cubic terms), characteristics of the student *S*, and whether or not the student had a TFA teacher in that year TFA_{iy}. Student characteristics include gender, race/ethnicity, age, parent's highest education level, whether the student ever had a disability, whether the student was ever identified as gifted, whether the student was over- or under-age for grade, and whether the student was ever retained.[xiii] In addition, the model includes fixed effects for school-grade-year combinations, π_{sgy}, and restricts the analysis sample to only school-grade-year combinations in which there was at least one TFA and one non-TFA teacher. This ensures that students who have TFA teachers are compared only

with other students who had the potential to have a TFA teacher but were assigned to a non-TFA teacher in the same school, grade, and year.

This approach addresses selection bias in several important ways, but also has limitations. It eliminates bias due to sorting of students into schools by comparing students within the same school. It also eliminates bias due to secular trends across years, including differences in reasons for having a TFA teacher one year and not the next, and bias due to differences across grades that might yield differences in achievement, such as the difficulty of the material. However, it is unable to address bias due to non-random sorting of students into classrooms within the same school and grade level that may be due to unobservables related to the student, the teacher, or parental preferences. Although the presence and magnitude of this bias in North Carolina and other settings is debated (Clotfelter et al., 2006; Koedel & Betts, 2011; Rothstein, 2009), this approach continues to be used in many administrative data applications for comparing the impact of TFA teachers on student achievement (Boyd et al., 2006; Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009). In all models, I cluster the standard errors at the classroom level to account for the non-independence of students in the same classrooms. Results are largely consistent in alternative models that cluster at the school and year level.

Initial models examine the relationship between having a TFA teacher and student test scores as described above. Subsequent models add controls for teacher qualifications to examine whether the main effects of TFA can be explained by observable teacher qualifications. I then examine whether the effect of TFA changes over time by examining the relationship between having a TFA teacher and student achievement in the early and later TFA periods (through 2005 and post-2005), and in models that examine these relationships in each year. I formally test for differences between the two periods using an interaction with TFA and the post-2005 period and

contrast a rapid subsequent improvement with a linear trend in TFA effects over time. Finally, to test whether or not teacher qualifications can explain any changes in the relationship between TFA and student achievement, I estimate models with terms for teacher qualifications, as well as the interaction of each of the teacher qualification measures and the post-2005 period.

The decision to isolate two distinct periods in TFA's history in North Carolina stems from policy changes made within the TFA organization during this time period that were intended to improve the effectiveness of TFA corps members. This analysis focuses on differences occurring before and after this particular point in time because of the organizational salience. It then contrasts this explanation with a more gradual pattern of change.

Results

Table 3 displays the main effects of TFA teachers on student achievement across school levels and subjects. The relationship between TFA and student achievement is positive in every subject and grade except elementary reading, although it differs in magnitude. TFA effects are by far the largest in high school science, where student achievement is 0.193 standard deviations higher on average in TFA than in non-TFA classrooms. The TFA effect in math is also fairly substantial at 0.133 standard deviations. Having a TFA teacher is also associated with significantly higher achievement in high school social studies, elementary and middle school math, and at marginally significant levels in high school English. The one exception to these positive associations between TFA and achievement is elementary reading, where the TFA coefficient is negative and not significant. Supplemental analyses examining each individual subject and grade level separately show that these individual results largely conform to the results presented in Table 3.[xiv]

[Insert Table 3 here]

Table 4 presents results from models examining whether the relationship between TFA teachers and student achievement is explained by observable teacher characteristics. Results examining models that add controls for the teacher's average Praxis certification scores, their years of experience, whether the teacher is a veteran, fully certified, has a master's degree or more, the selectivity of the teachers' undergraduate institution, and all qualification measures combined are shown in Table 4. This table presents only the coefficients for having a TFA teacher and the teacher qualification measures of interest (full results from these models are available upon request).

[Insert Table 4 here]

These results show that very few of the observable teacher characteristics account for the relationship between TFA and student achievement. Typically, including the teacher qualification measures either does not appreciably change the effect of TFA, or if anything, increases it, suggesting that students in TFA classrooms perform even better relative to students with non-TFA teachers who have similar qualifications. For example, once the control for whether the student has a veteran teacher is added to the models in Table 4, the coefficients for TFA remain significant and increase in every subject and grade level, sometimes by more than 50 percent. We see a similar pattern in models that control for all qualification measures simultaneously.

The one observable measure of teacher qualifications that does explain part of the relationship between TFA and student achievement for nearly all subjects is teacher Praxis scores. Including controls for average Praxis scores reduces the magnitude of most of the TFA coefficients by 9-37 percent. In addition, adding the Praxis control also increases the absolute magnitude of the significant, negative coefficient for Elementary reading, so that TFA teachers

have a negative effect of -0.023 standard deviations relative to teachers with similar average Praxis scores. While this difference is still relatively small, it is noteworthy, suggesting that TFA teachers would do even worse in reading if counterfactual teachers were stronger. Additionally, Barron's rankings do completely explain the effects of TFA on achievement in high school social studies and English, but not in other subjects, and net of these rankings, students in TFA classrooms perform worse on elementary reading.

Given previous findings that in some cases TFA teachers perform worse than veteran teachers, an additional robustness check compared TFA teachers to only veteran non-TFA teachers only. These models echo the results shown in Table 3, finding that TFA teachers outperform even veteran non-TFA teachers in all subjects except elementary reading, where their students score 0.025 standard deviations worse than students who have veteran teachers (results available upon request). Overall, the coefficients for the veteran teacher comparisons are of similar magnitude to those comparing TFA teachers to all non-TFA teachers.

Changes in TFA effects over time

I next examine whether TFA's association with student achievement changed as the TFA program matured. I begin by testing whether the effects of TFA vary across years. F-tests reject the null hypotheses that the effects are uniform across years. These F-tests are significant at conventional levels for seven of the eight subjects and marginally significant for the eighth (middle school reading), providing strong evidence that TFA's effects changed over time. In Table 5, I examine whether these changes coincide with the reforms introduced in 2005. The first columns for each subject replicate the regression results for the full 2000-2011 period shown in Table 3. Two additional columns per outcome show results restricted to the years 2000-2005 and 2006-2011.[xv]

[Insert Table 5 here]

The coefficients for elementary math, and middle school math and reading follow a similar pattern. In these three subjects, TFA teachers have a positive and significant association with achievement for the full panel, the coefficients for TFA are small and insignificant in the 2000-2005 period, and then the coefficients for TFA in the 2006-2011 period are significant and larger than the coefficients for the full panel. These results are indicative of an improvement in the effect of TFA from the earlier to the later period. In elementary reading, the results are only marginally significant in the early period, and indicate performance that is .084 standard deviations worse than in non-TFA classrooms. Overall, and in the later period, the association between TFA and elementary reading achievement is not significant, although the coefficients are still negative.

Models examining the earlier and later TFA periods for the four high school subjects also indicate improvements in the later period in three of the four subjects. In high school English, having a TFA teacher in the early period was associated with similar achievement in TFA and non-TFA classrooms, but by the later period, students in TFA classrooms outperformed those in non-TFA classrooms by 0.038 standard deviations. In high school science, TFA was positively associated with achievement in the earlier period and got even better as the TFA program matured, so that by the later period students of TFA science teachers had achievement that was more than two tenths of a standard deviation better than students in non-TFA science classrooms. In high school social studies, TFA achievement effects also improve substantially over time. Here TFA teachers were substantially worse than non-TFA teachers in the early period (0.42 standard deviations) and in the later period have improved to be more effective than non-TFA teachers by 0.072 standard deviations.[xvi] In high school math, the positive association between

having a TFA teacher and achievement was sustained across the period, but was slightly larger in the earlier period than the later period.

These patterns of improvement in the effects of TFA from an earlier to a later period suggest the possibility that changes occurred in response to the 2005 program updates. It is also possible that a small number of exceedingly successful years or slow, sustained improvement drives these results, rather than a rapid response to program updates. I estimate several additional models to interrogate exactly when and how rapidly these changes occurred. Figures 1 and 2 present results from models that examine the association between TFA and achievement in each school year and subject. The points represent the estimated TFA effect for the particular year and subject with lines indicating the 95 percent confidence intervals from regressions that adjust for other covariates and school-grade-year fixed effects, identical to the models described above.

[Insert Figures 1 and 2 here]

Each of the panels in Figure 1 show that the estimates get more stable over time as more TFA teachers taught in elementary and middle schools. Results also indicate small increases over time that fluctuate somewhat, rather than a discrete jump immediately following program reforms. Middle school math has the most robust pattern of improvement, with some positive and significant coefficients in years after 2005. However, even in this case there is little indication of a discrete, sustained jump beginning after 2005.

In contrast, Figure 2 suggests that in many years in high school, TFA students outperform non-TFA students in social studies, math, and science. In particular, TFA has positive effects on achievement in math and science most of the post-2005 years. This pattern is not sustained across all years, but it does suggest some sort of improvement after the reforms, though perhaps not beginning until a year or two after the reforms went into effect.

The results in Table 5 suggest that TFA effects improve over time in several subjects, but these models do not provide a formal test of whether or not TFA effects are significantly different across the two periods, which the annual results suggest might be the case in some subjects, but not others. Panel A of Table 6 reports results from models including an interaction term for TFA times post-2005 to test whether TFA effects are significantly different before and after the reforms. Results indicate significant differences over time in three subjects, middle school math, high school science, and high school social studies.[xvii]

While there are significant differences in these subjects between TFA in the early and late periods, this could result from gradual improvement over time rather than a discrete jump after the program changes. To test this, I estimate models (not shown) that include an interaction between TFA and post-2005 as well as an interaction for TFA and a continuous year variable. Among the subjects with a post-2005 increase, results indicate a mixed pattern. High school social studies most closely follows a pattern of a post-reform increase, although a small portion of the increase is also due to general improvement over time. Results for high school math indicate a general positive increase over time, and if anything a negative response to the reforms in the post-2005 period. The marginally significant increases in the post-2005 period in middle school math and high school science are not clearly attributable to either a post-reform jump or a gradual increase over time.

A test of the joint significance of the individual TFA by year interaction terms described in Figures 1 and 2 rejects the null hypothesis that the effects are uniform across years. This F-test is significant at conventional levels for seven of the eight subjects and marginally significant for the eighth (middle school reading). This provides evidence that TFA's effects were changing over time. This coupled with the limited signs of an immediate post-reform change suggest a

more gradual pattern of improvement rather than an immediate response to reform efforts.

[Insert Table 6 here]

Finally, Panels B-H of Table 6, summarize the key coefficients from models introducing measures of teacher qualification to test whether changes over time can be explained by differences in teacher qualification (full model results are available upon request). These results mirror those showing the relationship between the main effect of TFA and teacher qualification. The majority of the teacher qualification measures have little impact on the interaction of TFA and the post-2005 indicator. When I include all of the controls (Panel B), these qualifications account for roughly half of the positive changes in the association between TFA and achievement over time in high school science, and that the change across periods is no longer significant for high school science. By contrast, models examining the changes over time in middle school math and high school social studies are largely unchanged even when all of the qualification covariates are included.

Discussion

Teach For America has focused on improving educational opportunities and outcomes for marginalized for several decades. It does so by recruiting largely uncertified, new college graduates from elite institutions to teach in underserved schools. While this basic model of the program has not changed dramatically, the organization has evolved substantially in its training methods, organizational focus, and even where and whom it recruits. Few prior studies account for this evolution when considering potential changes in the relationship between TFA and student outcomes.

Previous work identified variation in the effect of TFA across subject areas, school levels, the distribution of student achievement and the match between the program and student

baseline proficiency (Author, 2016; Glazerman et al., 2006; Henry et al., 2010). I replicate these findings using statewide administrative data from North Carolina, and highlight that in many areas TFA's effects have been sustained or even increased through multiple evolutions of the program in ways that cannot be accounted for by observable measures of teacher qualifications.

Consistent with a growing number of studies, I find that TFA teachers outperform non-TFA teachers in most subjects, particularly at the secondary level. Even after programmatic changes overtime and large growth in corps size in North Carolina, TFA maintained its positive association with achievement in nearly every subject. Further, TFA teachers showed significant improvements in middle school math, and high school science and social studies.

While overall this study indicates that TFA's effects on achievement trend positive, results in two subject areas deviate from this pattern. The first is elementary reading, where the main effect of TFA is not significantly different from that of non-TFA teachers. When compared with veteran teachers, TFA's effects on elementary reading are negative. Some of the time-trend models do provide some slightly more positive news and suggest that TFA has improved from being significantly less effective in the earlier period to on par with the average non-TFA teacher in the later period. These findings suggest that the selectivity and training of TFA are still not sufficient to enable TFA teachers to outperform their non-TFA counterparts in elementary reading. This is largely consistent with previous experimental work on TFA (Author, 2016; Glazerman et al., 2006), but, see Clark et al.'s (2015) evidence on early grade TFA effects.

There is also a change in high school math over time that deviates from the trend of improvement seen in some other subjects. This is somewhat curious given the large positive main effects of TFA in high school math. The interaction coefficient for the later TFA period is negative, although this difference is not statistically significant. This result seems to be driven

particularly by a large negative effect of TFA in 2006, directly after major program updates, but TFA effects were again positive and significant in three of five subsequent years. Subject-specific analyses (not shown) suggest that this pattern was driven by declines in TFA's effects in Algebra I, suggesting that future work observing the classroom practices of TFA and non-TFA teachers would do well to examine Algebra classes in particular.

This study also examines the degree to which the TFA's achievement effects – and their change over time – can be attributed to differences in observable measures of teacher qualifications. I find evidence suggesting that most of the commonly used measures of observable teacher qualification cannot explain differences between TFA and non-TFA teachers overall. If anything, once measures like certification status or master's degrees are controlled for, TFA teachers perform even better relative to their similarly certified counterparts. This is perhaps not surprising, given that the vast majority of TFA corps members rank very poorly on traditional measures of teacher qualification (e.g. full certification, advanced degrees, experience, etc.). Thus, accounting for these differences only increases the relative achievement impacts of TFA teachers. Further, differences remain even when comparing TFA teachers to veteran teachers in the same schools. This is noteworthy, and given past criticisms of TFA centered on experience. Overall, these results suggest that something besides observable teacher qualifications is driving TFA's effects.

An alternative explanation for why students in TFA classrooms have higher achievement than those in non-TFA classrooms draws from the evolving nature of the TFA program itself. Results provide limited evidence that the changes stem from the overhaul of its training and coaching in the mid-2000s in most subjects. These improvements are not clearly an immediate response to program changes in the mid-2000s and in some cases seem to be more a result of

steady improvements over time. However, changes in the nature of the TFA program, whether due to selection procedures, matching of TFA teacher placements, or training might still account for some discrepant findings across prior TFA studies (Clark et al., 2013; Darling-Hammond et al., 2005; Henry et al., 2014; Laczko-Kerr & Berliner, 2002; Raymond & Fletcher, 2002; Xu et al., 2011).[xviii]

TFA has been criticized for its intensive focus on achievement, as well as some of its institutional structures and decisions that have supported this focus. While the literature examining the effects of TFA teachers on behavioral and non-academic outcomes is small, but growing, a weakness of this paper is its single focus on student achievement outcomes. However, given the dramatic changes the program has undergone over time, it is also important to investigate whether these changes have maintained the program's impact on achievement that others have identified. While examining similar trends in non-achievement outcomes would bolster this evidence, the data do not currently support this investigation. This evidence should be weighed among the multiple sources of evidence that districts and school leaders consider when assessing their teaching needs and interest in hiring TFA teachers.

In spite of these limitations, these results suggest that for many subjects, TFA teachers are associated with improved academic performance over more than a decade. This association has improved over time for several subjects. While the changes in teacher qualifications do not account for these improvements nor did changes immediately follow program reforms, positive trends provide suggestive evidence of gradual improvements over time. This is perhaps an indication that reforms took several years to make an appreciable impact on corps member practice. This pattern is not unprecedented. School improvement reforms in other areas also provide an example of reforms that take several years to yield substantial improvements (Author,

2017). Without historical data from the TFA organization, it is difficult to pinpoint the precise features of the training and selection evolution that contributed to TFA's improvements. However, given TFA's sustained presence in underperforming school districts and its increasingly positive impact on student achievement in many subjects, but not all, it merits continued investigation and critique. Such investigations should consider not only program impacts at any given time, but should consider how and why program impacts might have changed over time.

References

Author. (2014; 2016; 2017).

Backes, B., & Hansen, M. (2018). The impact of Teach For America on non-test academic outcomes. *Education Finance and Policy, 13(2)*, 168-193.

Backes, B., Hansen, M., Xu, Z., & Brady, V. (2016). Examining spillover effects from Teach For America corps members in Miami-Dade County public schools. *Journal of Teacher Education*, 0022487117752309.

Boyd, D. J., Dunlop, E., Lankford, H., Loeb, S., Mahler, P., O'Brien, R., & Wyckoff, J. (2012). *Alternative Certification in the Long Run: A Decade of Evidence on the Effects of Alternative Certification in New York City*. Presented at the 37th Annual Meeting of the Association of Education Policy and Finance, Boston, MA. Retrieved from aefpweb.org

Boyd, D. J., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2012). Recruiting effective math teachers: Evidence from New York City. *American Educational Research Journal*, *49*(6), 1008–1047.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 416–440.

Boyd, D. J., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1–2*, 176–216.

Boyd, D. J., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management*, *27*(4), 793–818.

Brewer, T. J., K. Kretchmar, B. Sondel, S. Ishmael, and M. McGlinn Manfra. 2016. "Teach For America's Preferential Treatment: School District Contracts, Hiring Decisions, and Employment Practices." *Education Policy Analysis Archives 24 (15)*: 1–32.

Brewer, T. J., and M. Wallis. 2015. "#TFA: The Intersection of Social Media and Education Reform." Critical Education 6 (13): 1–13.

Brooks, E., and K. Greene. 2013. "Problems, Politics, and Possibilities: Imagining a Teach For America That Really is for America." *Critical Education 4 (13)*: 68–92.

Carroll, S., Reichardt, R., Guarino, C., & Mejia, A. (2000). *The distribution of teachers among California's school districts and schools*. DTIC Document.

Cha, S.-H., & Cohen-Vogel, L. (2011). Why they quit: a focused look at teachers who leave for other occupations. *School Effectiveness and School Improvement*, *22*(4), 371–392.

Chiang, H. S., Clark, M. A., & McConnell, S. (2017). Supplying disadvantaged schools with effective teachers: Experimental evidence on secondary math teachers from Teach for America. *Journal of Policy Analysis and Management*, *36*(1), 97–125.

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

Clark, M. A., Chiang, H. S., Silva, T., McConnell, S., Sonnenfeld, K., Erbe, A., & Puma, M. (2013). *The effectiveness of secondary math teachers from Teach For America and the Teaching Fellows programs.* (No. NCEE 2013-4015). National Center for Education Evaluation and Regional Assistance.

Clark, M. A., Isenberg, E., Liu, A. Y., Makowsky, L., & Zukiewicz, M. (2015). *Assessing the Effectiveness of Teach For America's Investing in Innovation Scale-Up*. Princeton, NJ: Mathematica Policy Research.

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

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*(3), 655–681.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2015). The aftermath of accelerating algebra evidence from district policy initiatives. *Journal of Human Resources*, *50*(1), 159-188.

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Wheeler, J. (2006). High-poverty schools and the distribution of teachers and principals. *NCL Rev.*, *85*, 1345.

Cohen-Vogel, L., & Osborne-Lampkin, L. (2007). Allocating Quality: Collective Bargaining Agreements and Administrative Discretion over Teacher Assignment. *Educational Administration Quarterly*, *43*(4), 433–461.

Cohen-Vogel, L., Osborne-Lampkin, L., & Houck, E. (2013). New data, old patterns: The role of test scores in student assignment. In *The infrastructure of accountability: Mapping data use and its consequences* (pp. 141–161). Cambridge, MA: Harvard Education Press.

Corcoran, S. P., Evans, W. N., & Schwab, R. S. (2004). Changing labor market opportunities for women and the quality of teachers, 1957-2000. *American Economic Review*, *94*(2), 230–235.

Darling-Hammond, L. (1994). Who will speak for the children? *Phi Delta Kappan*, *76*(1), 21.

Darling-Hammond, L. (2003). Access to Quality Teaching: An Analysis of Inequality in California's Public Schools. *Santa Clara Law Review*, *43*(4), 1045.

Darling-Hammond, L., Brewer, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, *13*(42), 1–42.

Donaldson, M., Johnson, S. M. (2010). The price of misassignment: The role of teaching assignments in Teach for America teachers' exit from low-income schools and the teaching profession. Education Evaluation and Policy Analysis, 32, 299-323.

Esch, C. E., Chang-Ross, C. M., Guha, R., Humphrey, D. C., Shields, P. M., Tiffany-Morales, J. D., … Woodworth, K. R. (2005). The Status of the Teaching Profession, 2005. Full Report. *Center for the Future of Teaching and Learning*.

Farr, S. (2010). *Teaching as leadership: The highly effective teacher's guide to closing the achievement gap*. San Francisco, CA: Jossey-Bass.

Glazerman, S., Mayer, D., & Decker, P. (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.

Goldhaber, D., & Anthony, E. (2004). Can teacher quality be effectively assessed? *Unpublished Manuscript*.

Goldhaber, D., Gross, B., & Player, D. (2011). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management*, *30*(1), 57–87.

Goldhaber, D., & Hansen, M. (2010). Using performance on the job to inform teacher tenure decisions. *The American Economic Review*, *100*(2), 250–255.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources*, *39*(2), 326–354.

Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2014). The Effects of Teacher Entry Portals on Student Achievement.

Journal of Teacher Education, *65*(1), 7–23.

Henry, G. T., Thompson, C. T., Bastian, K. C., Fortner, C. K., Kershaw, D. C., Purtell, K. M., & Zulli, R. A. (2010). *Portal report: teacher preparation and student test scores in North Carolina*. Chapel Hill, NC: Carolina Institute for Public Policy, The University of North Carolina at Chapel Hill.

Irizarry, J., & Donaldson, M. L. (2012). Teach for America: The Latinization of US schools and the critical shortage of Latina/o teachers. *American Educational Research Journal, 49(1),* 155-194.

Jaramillo, M. (2002). *About Us: Recent press: Teach for America's Application Numbers Continue to Rise*. Teach For America. Retrieved from http://www.teachforamerica.org/pdfs/press/2002_applications_2003.pdf

Jennings, L. (2014, July 2). End-of-Year Updates From Teach For America - Los Angeles. Retrieved from http://www2.teachforamerica.org/webmail/29012/4870395/42528b2d03e95ec4ee4bc79f7fcb414d

Kane, T., Rockoff, J., & Staiger, D. O. (2008). What does teacher certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27*(6), 615–631.

Koedel, C., & Betts, J. R. (2011). Does student sorting invalidate value-added models of teacher effectiveness? An extended analysis of the Rothstein critique. *Education Finance and Policy, 6*, 18–42.

Kopp, W. (2011). *One day, all children* (3rd ed.). Cambridge, MA: Public Affairs, NY Perseus Book Group.

Kraft, M. A. (2017). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, 0916-8265R3.

Kretchmar, K., Sondel, B., & Ferrare, J. J. (2018). The power of the network: Teach for America's impact on the deregulation of teacher education. Educational Policy, 32(3), 423-453.

Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy. *Education Policy Analysis Archives, 10*(37). Retrieved from

http://epaa.asu.edu/epaa/v10n37/

Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy, 12(2),* 241-279.

Lahann, R., and E. Reagan. 2011. "Teach For America and the Politics of Progressive Neoliberalism." Teacher Education Quarterly 38 (1): 7–27.

McNew-Birren, J., Hildebrand, T., & Belknap, G. (2018). Strange bedfellows in science teacher preparation: conflicting perspectives on social justice presented in a Teach For America—university partnership. Cultural Studies of Science Education, 13(2), 437-462.

Mungal, A. S. 2016. "Teach For America, Relay Graduate School, and the Charter School Networks: The Making of a Parallel Education Structure." *Education Policy Analysis Archives 24 (17)*: 1–26.

NCERDC. (2012). *End of Course Test Documentation*. Durham, NC: Duke University.

ncpublicschools.org/fbs/accounting/data/. (2013). Data & Reports. Retrieved August 20, 2013, from http://www.ncpublicschools.org/fbs/accounting/data/

North Carolina Department of Public Schools. (2011, August 4). Evolution of the ABCs. Retrieved from http://www.ncpublicschools.org/docs/accountability/reporting/abc/2010-11/abcevolution.pdf

Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvemet. *Journal of Public Economics, 130,* 105-119.

Raymond, M., & Fletcher, S. (2002). The Teach for America Evaluation. *Education Next*, *2*(1), 7.

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education Finance and Policy*, *4*(4), 537–571.

Sawchuk, S. (2009). Growth model. *Education Week*, *29*(3), 27–29.

Scafidi, B., Sjoquist, D. L., & Stinebrickner, T. R. (2006). Do Teachers Really Leave for Higher Paying Jobs in Alternative Occupations? *Advances in Economic Analysis & Policy*, *6*(1).

Schoeneberger, J. A., Dever, K. A., & Tingle, L. (2009). *Evaluation of Teach For America in Charlotte-Mecklenburg Schools*. Charlotte, NC: Center for Research and Evaluation Office of Accountability: Charlotte-Mecklenburg Schools.

Sondel, B. (2015). Raising citizens or raising test scores? Teach for America, "no excuses"

charters, and the development of the neoliberal citizen. *Theory & Research in Social Education, 43*, 289–313.

Teach For America. (2017). Where We Work. Retrieved January 16, 2017, from https://www.teachforamerica.org/where-we-work

teachforamerica.org. (2013a). Our Organization. Retrieved August 20, 2013, from http://www.teachforamerica.org/our-organization

teachforamerica.org. (2013b). Where We Work. Retrieved January 22, 2013, from http://www.teachforamerica.org/where-we-work

teachforamerica.org. (2013c, August 6). About Teach For America 2013-14. Retrieved from https://www.teachforamerica.org/sites/default/files/2013-14_press_kit_updated_08_6_13.pdf

teachforamerica.org. (2016, January 29). The First 25 Years | One Day Magazine. Retrieved January 14, 2017, from https://www.teachforamerica.org/one-day-magazine/first-25-years.

Tourangeau, M. (2003, January). Teach for America. Retrieved June 30, 2014, from http://learningtogive.org/papers/paper161.html

Trujillo, T., Scott, J., & Rivera, M. (2017). Follow the yellow brick road: Teach for America and the making of educational leaders. *American Journal of Education, 123(3),* 353-391.

Vance, V. S., & Schlechty, P. C. (1982). The distribution of academic ability in the teaching force: Policy implications. *The Phi Delta Kappan*, *64*(1), 22–27.

Ware, A., LaTurner, R. J., Parsons, J., Okulicz-Kozaryn, A., Garland, M., & Klopfenstein, K. (2011). *Teacher preparation programs and Teach for America research study*. Education Research Center: University of Texas at Dallas.

White, T. C. (2016). Teach For America's paradoxical diversity initiative: Race, policy, and Black teacher displacement in urban schools. *Education Policy Analysis Archives, 24*, 16.

Xu, Z., Hannaway, J., & Taylor, C. (2011). Making a difference? The effects of Teach For America in high school. *Journal of Policy Analysis and Management, 30*(3), 447–469.

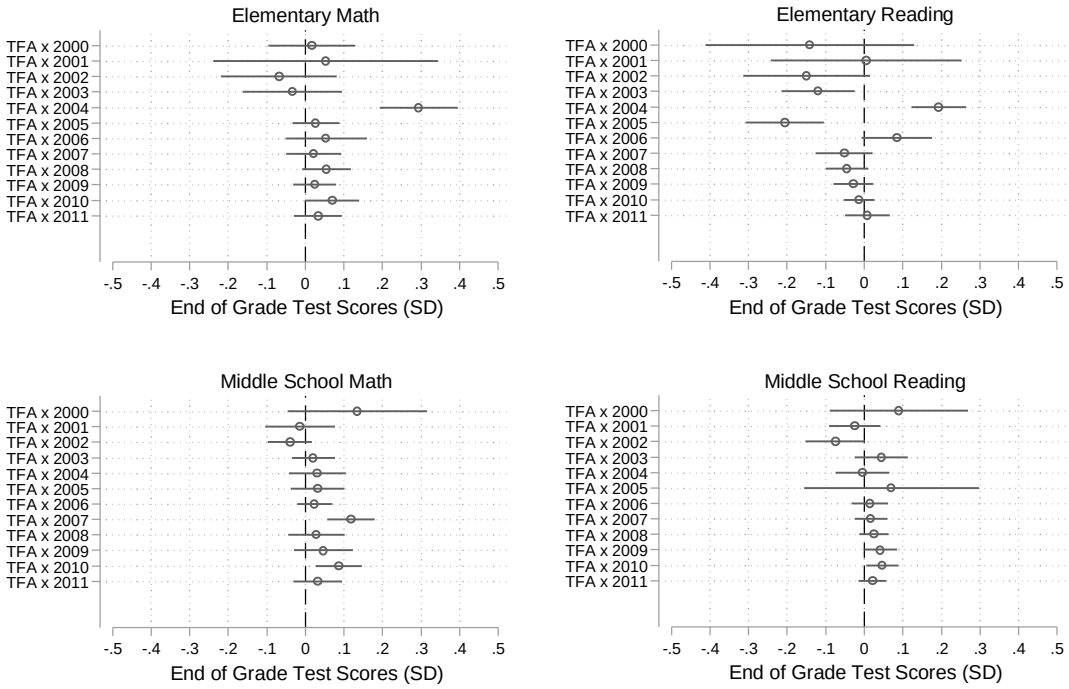

Figure 1. TFA effects by subject and year for grades four through five and six through eight. Coefficients presented represent the sum of the main effect of TFA and the interaction between having a TFA teacher and an indicator for each school year. Models include school-grade-year fixed effects, controls for lagged math and reading scores, including squared and cubic terms for both test scores, and student demographics. 95 percent confidence intervals reported for each regression coefficient. Standard errors clustered at the classroom level.

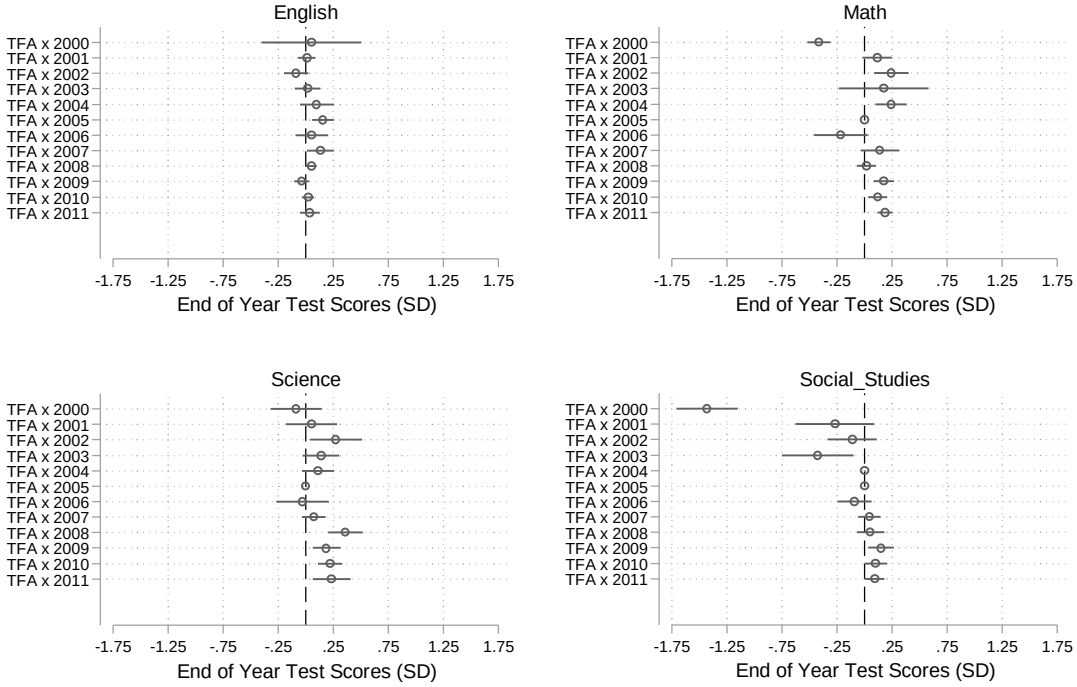

Figure 2. TFA effects by subject and year for grades nine through twelve. Coefficients presented represent the sum of the main effect of TFA and the interaction between having a TFA teacher and an indicator for each school year. Models include school-subject-year fixed effects, controls for lagged math and reading scores, including squared and cubic terms for both test scores, and student demographics. 95 percent confidence intervals reported for each regression coefficient. Standard errors clustered at the classroom level.

Table 1

Number of TFA teachers placed by region and year, and subject and year, grades K-12

	Region		***Subject Area***						
Year	Charlotte	Eastern North Carolina	Grades 3-8	HS English	HS Math	HS Science	HS Social Studies	Untested Subject	Total
1999-2000	0	28	15	4	1	6	2	0	28
2000-2001	0	25	14	4	1	5	1	0	25
2001-2002	0	26	16	1	1	8	0	0	26
2002-2003	0	41	29	6	2	2	2	0	41
2003-2004	0	7	3	1	1	2	0	0	7
2004-2005	0	1	1	0	0	0	0	0	1
2005-2006	54	43	78	6	2	1	5	5	97
2006-2007	58	60	60	6	5	8	10	29	118
2007-2008	55	51	59	3	6	6	2	30	106
2008-2009	98	82	93	5	7	17	7	51	180
2009-2010	63	56	51	3	13	6	5	41	119
2010-2011	104	52	75	8	11	5	5	52	156
Total	432	472	494	47	50	66	39	208	904

Table 2

Selected student and teacher characteristics across comparison samples

	TFA Classrooms	Non-TFA, Same School-Grade-Year	Non-TFA, Same School, Different Grade/Yr	Non-TFA, Same LEA, Different School	Non-TFA, Different LEA	Total
Student Characteristics						
Parent with college degree or higher	23.7	35.3	48	42	34.8	41
Ever identified as LEP	10.8	11	8.4	6.4	7.2	7.2
Student Ethnicity						
White	15.6	28.7	46.1	58.1	70.7	57.6
Black	67.5	53.4	39.5	27.7	17.5	28.6
Hispanic	10.6	10.8	8	6.9	7	7.3
Has some type of disability	15.6	15.5	15.6	16.6	16.9	16.4
Ever identified as gifted	10.3	16.3	21.5	22.3	19.5	21.1
Over-age for grade	31.5	29.3	23.1	22.1	21.9	22.6
Ever repeated grade	3.9	3.6	2.5	1.8	1.5	1.9
Standardized Lagged Test Scores (gr 4-8)						
Math - Mean	-0.31	-0.2	-0.03	0.03	0	0.01
Math - SD	0.86	0.91	0.88	0.82	0.81	0.83
Reading - Mean	-0.31	-0.21	-0.04	0.03	0.01	0
Reading - SD	0.87	0.92	0.87	0.81	0.82	0.83
Standardized 8th Grade Test Scores						
Math - Mean	-0.36	-0.13	0.13	0.15	0.15	0.13
Math - SD	0.86	0.97	0.98	0.96	0.93	0.96
Reading - Mean	-0.39	-0.14	0.14	0.15	0.13	0.12
Reading - SD	0.91	0.99	0.95	0.93	0.92	0.94
Teacher Characteristics						
Mean Years of Experience	0.9	11.4	12.2	12.8	15.1	13
% Veteran (>3 Years of exp)	3.3	48.5	47.6	48.3	34.5	44.3
% Fully Certified	26.5	77.6	74.6	79.7	76.5	77.2
% with an MA or more	4.7	31.2	25.8	26.4	25.6	26.1
Praxis Certification Score (std)	0.5	0.1	0.1	0.2	0.2	0.2
N (Students)	116715	1150219	5427891	14753375	7046688	28494888
N (Teachers)	904	4,741	36,198	89,832	42,350	174,025

Administrative data from North Carolina provided by the NCERDC; Test scores are standardized within subject, grade, and year to have a mean of zero and a standard deviation of one and are then shown for different subsamples of North Carolina students.

Table 3

Impact of TFA on achievement in elementary, middle, & high school

	Elementary School Math	Elementary School Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
TFA	**0.041****	-0.019	**0.048*****	**0.024****	**0.035+**	**0.133*****	**0.193*****	**0.054***
	(0.014)	(0.012)	(0.011)	(0.008)	(0.018)	(0.021)	(0.028)	(0.025)
School-grade-year/ subject-grade-year FEs	X	X	X	X	X	X	X	X
Lagged math & reading scores	X	X	X	X	X	X	X	X
Student covariates	X	X	X	X	X	X	X	X
R-squared	0.601	0.566	0.720	0.667	0.652	0.563	0.498	0.528
Observations	520504	516370	1046862	1046551	317953	675705	541170	476165

+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001; C23Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors clustered at the classroom level in parentheses. Student controls include: student race/ethnicity, if the student was ever identified as Limited English Proficient, if the student ever had an identified disability, if they were ever identified as gifted and talented, if they were identified as under- or over-age for grade, their parent's education level, lagged or eighth grade math and reading scores including squared and cubic terms, and indicators for missing values on any of these controls.

Table 4

Impact of TFA on achievement, elementary, middle, and high school; controlling for teacher education and experience

	Elementary School Math	Elementary School Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Studies
Panel A: Baseline Model								
TFA	**0.041****	-0.019	**0.048*****	**0.024****	**0.035+**	**0.133*****	**0.193*****	**0.054***
	(0.014)	(0.012)	(0.011)	(0.008)	(0.018)	(0.021)	(0.028)	(0.025)
Panel B: All Education & Experience Variables Simultaneously								
TFA	**0.050****	-0.002	**0.070*****	**0.037*****	**0.043***	**0.151*****	**0.234*****	**0.084****
	(0.015)	(0.013)	(0.012)	(0.008)	(0.020)	(0.022)	(0.031)	(0.027)
Panel C: Praxis Scores								
TFA	**0.027+**	**-0.023***	**0.040*****	**0.021****	0.014	**0.116*****	**0.179*****	0.034
	(0.014)	(0.012)	(0.011)	(0.008)	(0.018)	(0.021)	(0.028)	(0.025)
Mean Praxis Score	**0.019*****	**0.005****	**0.016*****	**0.005*****	**0.035*****	**0.012*****	**0.025*****	**0.035*****
	(0.002)	(0.002)	(0.002)	(0.001)	(0.004)	(0.004)	(0.005)	(0.005)
Panel D: Years of Experience								
TFA	**0.050*****	-0.009	**0.057*****	**0.026****	**0.044***	**0.147*****	**0.203*****	**0.064***
	(0.014)	(0.012)	(0.011)	(0.008)	(0.019)	(0.020)	(0.028)	(0.025)
Years of Experience	**0.001*****	**0.002*****	**0.001*****	**0.001*****	**0.002*****	**0.002*****	0.000	**0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Panel E: Veteran Teachers								
TFA	**0.058*****	-0.001	**0.060*****	**0.029*****	**0.046***	**0.160*****	**0.216*****	**0.079****
	(0.014)	(0.012)	(0.012)	(0.008)	(0.019)	(0.021)	(0.028)	(0.025)
Veteran Teacher (>3 yrs exper)	**0.040*****	**0.043*****	**0.028*****	**0.013*****	**0.053*****	**0.070*****	**0.053*****	**0.051*****
	(0.004)	(0.004)	(0.004)	(0.003)	(0.008)	(0.009)	(0.012)	(0.010)
Panel F: Fully Certified								
TFA	**0.069*****	0.001	**0.071*****	**0.042*****	**0.066*****	**0.187*****	**0.264*****	**0.110*****
	(0.015)	(0.012)	(0.012)	(0.008)	(0.018)	(0.021)	(0.030)	(0.027)
Fully Certified	**0.054*****	**0.038*****	**0.046*****	**0.040*****	**0.104*****	**0.122*****	**0.134*****	**0.124*****
	(0.010)	(0.009)	(0.005)	(0.005)	(0.013)	(0.011)	(0.016)	(0.018)
Panel G: Has an MA or more								
TFA	**0.041****	-0.017	**0.051*****	**0.023****	**0.032+**	**0.128*****	**0.191*****	**0.055***
	(0.014)	(0.012)	(0.011)	(0.008)	(0.018)	(0.021)	(0.028)	(0.025)
Has MA or more	**0.007***	**0.009****	**0.011*****	0.001	-0.003	0.005	0.000	0.012
	(0.004)	(0.003)	(0.003)	(0.002)	(0.007)	(0.007)	(0.009)	(0.008)
Panel H: Barron's Rankings								
TFA	**0.026+**	**-0.025***	**0.042*****	**0.019***	0.021	**0.100*****	**0.170*****	0.029
	(0.014)	(0.012)	(0.012)	(0.008)	(0.019)	(0.021)	(0.029)	(0.025)
Very Competitive Plus	**-0.019+**	-0.007	-0.009	-0.001	**0.044***	-0.019	**0.060****	0.005
	(0.010)	(0.010)	(0.008)	(0.006)	(0.022)	(0.022)	(0.022)	(0.021)
Highly Competitive	**0.025***	0.013	**0.014+**	**0.012+**	-0.001	0.003	**0.038+**	**0.077*****
	(0.010)	(0.009)	(0.008)	(0.007)	(0.016)	(0.014)	(0.020)	(0.020)
Highly Competitive Plus	0.015	0.008	**0.024***	**0.019***	0.021	0.019	0.014	0.002
	(0.017)	(0.015)	(0.011)	(0.009)	(0.024)	(0.019)	(0.030)	(0.022)
Most Competitive	**0.043*****	**0.017***	0.006	-0.002	**0.026****	**0.057*****	**0.043****	**0.083*****
	(0.009)	(0.008)	(0.005)	(0.004)	(0.009)	(0.011)	(0.014)	(0.012)

+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less, unranked BA institution; Standard errors clustered at the classroom level in parentheses. Models include fixed effects for school-grade-year or school-subject-year combinations and lagged or eighth grade test scores with squared and cubed terms.

Table 5

Impact of TFA on same-year tests in elementary, middle, and high school, by early and established TFA

	Elementary School Math			Elementary School Reading		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
TFA	**0.041****	0.011	**0.041****	-0.019	**-0.084+**	-0.015
	(0.014)	(0.042)	(0.014)	(0.012)	(0.044)	(0.012)
R-squared	0.601	0.613	0.593	0.566	0.573	0.563
Observations	520504	249449	271055	516370	247565	268805
	Middle School Math			**Middle School Reading**		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
TFA	**0.048*****	0.017	**0.055*****	**0.024****	0.006	**0.027****
	(0.011)	(0.017)	(0.014)	(0.008)	(0.019)	(0.008)
R-squared	0.720	0.711	0.734	0.667	0.648	0.690
Observations	1046862	501363	545499	1046551	499760	546791
	High School English			**High School Math**		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
TFA	**0.035+**	0.028	**0.038***	**0.133*****	**0.212****	**0.124*****
	(0.018)	(0.044)	(0.019)	(0.021)	(0.066)	(0.022)
R-squared	0.654	0.632	0.681	0.564	0.559	0.573
Observations	312920	143768	169152	667394	312702	354692
	High School Science			**High School Social Studies**		
	2001-2011	**2001-2005**	**2006-2011**	**2001-2011**	**2001-2005**	**2006-2011**
TFA	**0.193*****	**0.114***	**0.215*****	**0.054***	**-0.423****	**0.072****
	(0.028)	(0.050)	(0.033)	(0.025)	(0.131)	(0.025)
R-squared	0.500	0.490	0.522	0.528	0.497	0.565
Observations	534255	280795	253460	466059	166336	299723

+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Standard errors clustered at the classroom level in parentheses. Models include school-grade-year or school-subject-year fixed effects. Student controls inlcude: student race/ethnicity, if the student was ever identified as Limited English Proficient, if the student ever had an identified disability, if they were ever identified as gifted and talented, if they were identified as under- or over-age for grade, their parent's education level, lagged test scores or eighth grade test scores including squared and cubed terms, and indicators for missing values on any of these controls.

Table 6

Impact of TFA in elementary, middle, and high school; over time and net of teacher education and experience

	Elementary School Math	Elementary School Reading	Middle School Math	Middle School Reading	High School English	High School Math	High School Science	High School Social Stds.
Panel A: Baseline Model								
TFA	0.014	**-0.083+**	0.018	0.009	0.020	**0.201****	**0.115***	**-0.510*****
	(0.042)	(0.044)	(0.018)	(0.020)	(0.048)	(0.066)	(0.052)	(0.149)
TFA*Post-2005	0.028	0.069	**0.039+**	0.019	0.020	-0.073	**0.108 +**	**0.586*****
	(0.045)	(0.046)	(0.023)	(0.021)	(0.052)	(0.069)	(0.061)	(0.151)
Panel B: Net of All Education & Experience Controls								
TFA	0.050	-0.049	**0.037+**	0.022	0.062	**0.256*****	**0.204*****	**-0.386****
	(0.041)	(0.047)	(0.020)	(0.020)	(0.051)	(0.070)	(0.055)	(0.148)
TFA*Post-2005	0.000	0.050	**0.042+**	0.018	-0.024	-0.115	0.043	**0.486****
	(0.044)	(0.048)	(0.024)	(0.022)	(0.054)	(0.073)	(0.063)	(0.150)
Panel C: Net of Praxis Scores								
TFA	0.018	-0.080	0.018	0.009	0.013	**0.221*****	**0.120***	**-0.503*****
	(0.041)	(0.044)	(0.018)	(0.020)	(0.049)	(0.066)	(0.053)	(0.149)
TFA*Post-2005	0.008	0.062	0.028	0.016	0.011	-0.113	0.090	**0.563*****
	(0.044)	(0.046)	(0.023)	(0.021)	(0.052)	(0.069)	(0.062)	(0.152)
Panel D: Net of Years of Experience								
TFA	0.025	-0.074	0.024	0.010	0.040	**0.223*****	**0.139****	**-0.478*****
	(0.042)	(0.044)	(0.018)	(0.019)	(0.049)	(0.064)	(0.051)	(0.145)
TFA*Post-2005	0.026	0.070	0.043	0.019	0.000	-0.090	0.083	**0.555*****
	(0.044)	(0.046)	(0.023)	(0.021)	(0.053)	(0.068)	(0.061)	(0.147)
Panel E: Net of Veteran Teacher Indicator (>3 yrs. expr.)								
TFA	0.032	-0.066	0.028	0.011	0.044	**0.226*****	**0.151****	**-0.464****
	(0.042)	(0.044)	(0.018)	(0.019)	(0.049)	(0.064)	(0.050)	(0.147)
TFA*Post-2005	0.025	0.068	0.041	0.021	0.000	-0.083	0.080	**0.556*****
	(0.044)	(0.046)	(0.023)	(0.021)	(0.052)	(0.067)	(0.060)	(0.149)
Panel F: Net of Fully Certification Indicator								
TFA	0.046	-0.062	**0.042***	0.032	**0.089+**	**0.254*****	**0.209*****	**-0.400****
	(0.043)	(0.045)	(0.019)	(0.020)	(0.048)	(0.067)	(0.053)	(0.151)
TFA*Post-2005	0.020	0.066	0.039	0.012	-0.032	-0.085	0.075	**0.528*****
	(0.045)	(0.046)	(0.023)	(0.022)	(0.051)	(0.070)	(0.062)	(0.152)
Panel G: Net of MA or Higher Education Indicator								
TFA	0.007	**-0.086***	0.017	0.007	0.008	**0.163***	**0.111***	**-0.514*****
	(0.042)	(0.044)	(0.018)	(0.020)	(0.048)	(0.065)	(0.052)	(0.149)
TFA*Post-2005	0.036	0.073	0.044	0.019	0.031	-0.038	**0.110+**	**0.593*****
	(0.044)	(0.046)	(0.023)	(0.021)	(0.052)	(0.069)	(0.061)	(0.151)
Panel H: Net of Barron's Selectivity Rankings								
TFA	-0.006	**-0.099***	0.013	0.006	-0.015	0.114	0.088	**-0.530*****
	(0.043)	(0.045)	(0.018)	(0.020)	(0.052)	(0.072)	(0.056)	(0.151)
TFA*Post-2005	0.035	0.082	0.039	0.018	0.056	-0.002	**0.122+**	**0.582*****
	(0.046)	(0.046)	(0.023)	(0.022)	(0.055)	(0.075)	(0.065)	(0.153)

+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001; Omitted categories are: White; Parent Ed. High Sch. or less; Result present coefficinets from individual models. Standard errors clustered at the classroom level in parentheses. Models include fixed effects for school-grade-year or school-subject-year combinations.

[i] Active corps members do not exceed 3 percent of teachers in the placement districts with the largest numbers of TFA corps members. In most districts, they represent only one percent of teachers or fewer. Including alumni in the calculation does increase their presence to as much at 12 percent in heavily saturated districts like Houston, but it is not clear if all of the alumni identified by TFA in a given region are still classroom teachers (teachforamerica.org, 2013b).

[ii] Matched records are available for all students in North Carolina in grades K through 12 for twelve cohorts of students from the years 1999/2000 through 2010/2011. In a given school year, approximately 100,000 teachers educate roughly 1.4 million students in grades K-12 (ncpublicschools.org/fbs/accounting/data/, 2013). Analyses for this study concentrate on the students in grades 3-12 because no standardized testing occurs in North Carolina before grade 3.

While in most cases, the classroom teacher of record was the test administrator, in some cases, the test was administered by another adult in the school (NCERDC, 2012). The NCERDC provides guidelines for researchers to evaluate the likelihood that the test administrator is the classroom teacher, and I follow Goldhaber and Hansen (2010) in restricting my sample to teachers that were able to be matched to other district personnel records. For years 2006 and later, I use course code matches to assist in selecting the appropriate teacher and course match. These course codes are not available for years before 2006.

In elementary school, Goldhaber and Anthony (2004) contacted state officials who suggested that 90 percent of the time, the test administrator and the classroom teacher are the same individual. They then contacted administrators in the 20 largest districts and found that this match rate was 80 percent. In high school, Clotfelter, Ladd, and Vigdor (2010) link classroom data to student data using the classroom instructor code and a student exam proctor code. They then verify this match based on a fit statistic using student demographics, finding a match rate of 70-75 percent. Xu et al. (2011) also use a similar matching and verification method, matching approximately 84 percent of their students to teachers.

[iii] TFA identifies corps member years using the fall of their entering year, while I refer to school years using their fall and spring years. Thus the TFA teachers in this sample taught from the 1999/2000 school year through the 2010/2011 school year.

[iv] LEAs, or Local Education Agencies, are analogous to school districts in other state contexts. They include unique labels for public charter schools located within the boundaries of other LEAs. In 2012/2013 there were 115 LEAs and 107 charter schools (http://www.ncpublicschools.org/docs/fbs/resources/leacharterlist.pdf, 2013).

[v] The NCERDC made several attempts at matching based on different combinations of the provided demographic characteristics, including matches based on similar names (e.g. nicknames: Margaret/Maggie; last names: Duncan/Duncann), similar SSNs (i.e., 123456789/124356789), and allowing last names and school names to vary from those provided by TFA. A small portion of those successfully matched (7 percent) were matched on SSN alone. To maximize the total number of TFA teachers, all of those successfully matched, even those matched on SSN alone, are included in the analysis as TFA teachers.

TFA originally provided 1566 observations, however, 64 had repeated identifying information and were considered repeat observations. 87 percent of the unique observations provided by TFA were matched by NCERDC. Of the 198 observations that were unsuccessfully matched, most did not end up participating in TFA because they declined the offer, did not show up at Institute, resigned from TFA, or were released from the program due to an emergency. Only 36 of the unmatched observations were indicated as program completers by TFA (18 percent of all of the unmatched observations, 3 percent of the unique observations provided by TFA). The remaining teachers not assigned to a tested subject were assigned to grades K-2 or to other subjects in grades 6-12, such as foreign languages or middle school social studies. As a result, these teachers are not included in the current analysis sample, nor the descriptive information included in Table 2 and Appendix Table 1.

[vi] It is important to note that Praxis test scores are standardized within examination year and subject before being averaged within person. The sample for which these scores are measured includes teachers that instruct untested subjects, which have lower average Praxis scores than those teaching tested subjects.

[vii] Over 90 percent of students change schools between 5th and 6th grade, compared with roughly 30 percent in other grades, suggesting that the majority of 6th grades in North Carolina are part of middle schools, and thus I combine grades 6-8 in my analyses.

[viii] As Clotfelter, Ladd, and Vigdor (2015) describe, there is some algebra acceleration in North Carolina during this period where students in middle school grades also take some high school assessments. Across the panel, roughly 25,000 students in middle school grades take the Algebra 1 assessment. Given that the focus of this paper is comparing students within the same schools, grades, and years who do and do not have TFA teachers, I exclude the students in non-high school grades who take these assessments from my analyses and only focus on high school students who took these tests. This may understate the impact of TFA teachers on students who are on accelerated trajectories in these middle school grades, but such students are less likely to attend schools with TFA teachers in the first place.

[ix] If students had multiple test scores within a given year, I used the mean value of scores within the valid range for each subject and year before standardizing.

[x] In cases where students' lagged or eighth grade test scores were missing, I included an indicator for missing data to prevent the observations from dropping out of the model.

[xi] Being over- and under-age for grade is identified using the student's modal birth date across all observations, which is then differenced from the birthday cutoff date for the state of North Carolina, which is September 1. To protect the privacy of the students in the data, all birthdates were set to the 15th of the month by the NCERDC, so as a result over- and under-age cutoffs are both inclusive of the month of September and thus students are considered neither under- nor over-age for a 13-month window rather than a 12-month window.

[xii] An alternative approach, which estimates teacher-specific fixed effects and compares the estimates of different groups of teachers, is also frequently used in the literature examining the impact of teacher on student achievement. This approach is less preferred here because such value-added estimates are noisier with smaller numbers of students for each teacher. This difficulty is often overcome by using multiple years of data for each teacher with several different classrooms of students. Given that TFA teachers are largely first and second year teachers and do not have very many years of students to estimate from, they have noisier estimates using the teacher fixed effects approach than more experienced teachers. Nonetheless, this approach is used by some to estimate value-added estimates for TFA teachers (e.g., Henry et al., 2010).

[xiii] Supplemental models without controls for student characteristics yield similar results, suggesting that bias due to including endogenous control variables is not driving my results.

[xiv] The one exception is social studies, where I found large negative effects in Economics, Legal, and Political Systems (-0.971 SD), and moderate positive effects in the other two subjects. It is important to note that the exams in Social Studies changed more substantially than in other subjects. The ELP exam was given from before 2000 until 2004. The U.S. History exam was given throughout the years of this study, with the exception of 2005, and the Civics and Economics exam, which replaced much of the content of the ELP exam, was first given in 2006 and continued through 2011 and afterwards (North Carolina Department of Public Schools, 2011). Importantly, however, the changes in the tests occurred for all students, and thus affected both students with and without TFA teachers equally.

[xv] The primary models include TFA alumni, as these individuals are still considered TFA-trained teachers even after completion of the two year program. Supplementary models that exclude alumni are largely consistent, with the exception of some of the high school subjects in the early period in which there are some differences including two subjects in which there are too few TFA teachers to estimate TFA effects.

[xvi] Given the changes in the tested subjects in social studies over time, the social studies results should be judged with a greater degree of caution than results from the other subject areas and grade levels.

[xvii] Supplemental models comparing TFA teachers to only veteran non-TFA teachers suggest that the TFA improvement occurs in these subjects regardless of the experience level of the counterfactual teachers.

[xviii] Differences between the earlier and later studies might also be due in part to differences in research design, though it is worth noting that if these differences were only due to differences in the research design it seems unlikely that the present study would have found differences over time.